# PAX6 protein-protein interactions.

Simon T. Cooper

Ph.D.

The University of Edinburgh

December 2003

# Declaration

I declare that this thesis has been composed by myself, that all the work is my own unless otherwise clearly stated and that it has not been submitted for any other degree or professional qualification.

Simon T. Cooper

December 2003

# Acknowledgements

First and foremost I would like to thank my sanity for holding out long enough for me to complete this thesis and the lab work that preceded it – it was a close call. Next and equally as amazing is the help I have received from my supervisor Dr Isabel Hanson. Not only has she been the font of all knowledge but also she has been a source of inspiration and encouragement and has tirelessly proofread umpteen chapters of my thesis – my deepest thanks to you Isabel.

I would also like to mention Alison Brown who has been my lab guardian and reagent pixie, many thanks to you not only for your support throughout my PhD in the lab but also those magic hands that helped me train for, if not run in, the Edinburgh Marathon.

My thanks also to the PhD students old and new, who have helped in all aspects of my time at the MMC. Special thanks must go to my long suffering partners in crime, Jean O'Donoghue and Miss Sarah Warne as she was and Mrs Sarah Underwood as she is now, who have been with me since the beginning and through to the end.

I would also like to thank Rachel James for her antibody and everyone else within the Molecular Medicine Centre who has helped me throughout my time in the lab. I thank you for supporting me with my crazy ideas such as walking across the desert and I thank you for your tolerance of my slightly silly sense of humour. I would also like to thank Anne Seawright, Kathy Williamson and Veronica van Heyningen for their kind donation of tissue culture cells and patient RNA samples.

Outside the MMC, (is there really such a place), I would like to thank all my friends for being there to distract me when I needed it, but especially: to Jon and Gillian for keeping me fed and watered, I promise I will cook for you guys one of these day; to Bizz and Ed for helping me satisfy my Club sandwich obsession; to KT, Jill and Bizz again, for my culture outings; to Mr Edie who, it must be said, didn't do much but how could I leave the buffoon out; and to Dave and Kirsty who just make me laugh and keep me happy.

Finally to my family Mum, Dad, Matt and Paula – you are all very special to me, there'll be no thanks for you, but you are all lovely.

# Abstract

The gene *PAX6* is located on chromosome 11 (11p13) and encodes a transcription factor (PAX6) that is expressed early in development. The PAX6 protein is expressed in the developing eye, regions of the brain, central nervous system (CNS), nasal epithelium and pancreas. PAX6 is best known for its role in eye development with heterozygous mutations causing congenital ocular malformations. However, it must be remembered that PAX6 has multiple functions in the brain including specification of neuronal subtypes and axon guidance.

There is growing understanding of the role of PAX6 as a transcription factor during development, and many of its DNA targets have recently been defined. However, almost nothing is known about the proteins with which PAX6 interacts.

In the initial stage of my research I identified a conserved region consisting of the final 32 amino acids of the PST (proline, serine and threonine rich) domain of PAX6. Based on sequence homology and secondary structure predictions I classed this region as a novel domain, the 'C terminal domain'. Next I used the yeast 2-hybrid system to investigate possible PAX6 protein interactions. By screening a mouse brain cDNA library with the C terminal domain and whole PST domain, I identified three novel and interesting interactors, Homer3, Dncl1 and Trim11.

I re-confirmed these interactions in a pairwise manner using the yeast 2-hybrid system, and I showed that the C terminal domain was vital for the interactions between PAX6 and Homer3 or Dncl1. Furthermore, certain C terminal mutations that are known to cause ocular malformations in patients are also sufficient to reduce or abolish these interactions. I attempted to further characterise the interactions by co-immunoprecipitation. However, this was not possible due to technical difficulties.

Although speculative at present, the finding that PAX6 may interact with Homer3 (a component of the post synaptic density) and Dncl1 (a subunit of the motor protein dynein) alludes to an interesting new role for PAX6 in neurogenesis.

# List of Contents

# List of Figures

# List of Tables

# Chapter 1 Introduction

## 1.1 *PAX6* – A brief gene outline.

*PAX6* has been identified as a master control gene for eye development and is located on chromosome 11 (11p13). The gene comprises 15 exons that span 28kb and this is illustrated in Figure 1.1 (Glaser *et al.*, 1992; Xu *et al.*, 1999b). The *PAX6* gene was first identified by positional cloning as a candidate aniridia gene (Ton *et al.*, 1991). *PAX6* is expressed early in development not only in the eye, but also in regions of the brain, the developing central nervous system, nasal epithelium and the pancreas (Prosser and van Heyningen, 1998).

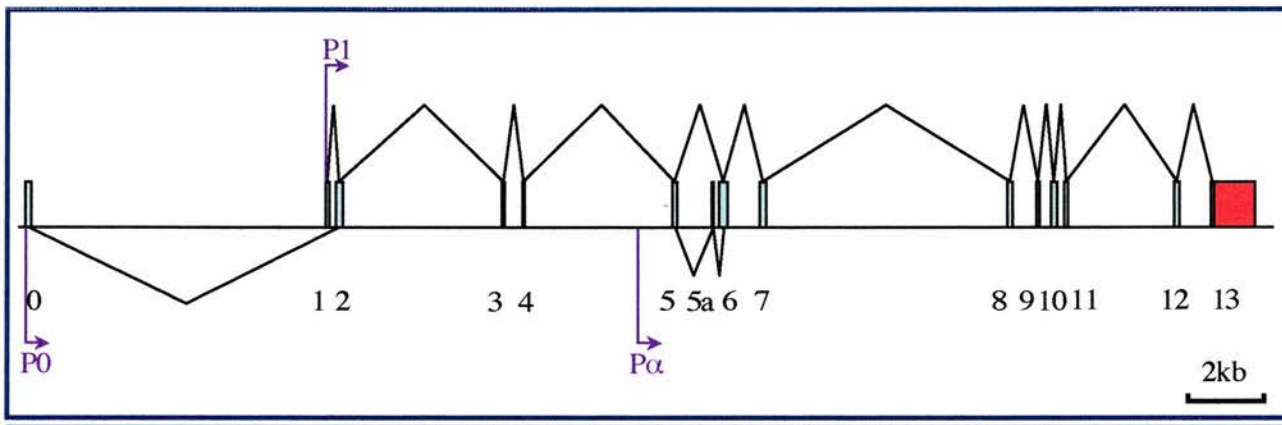



**Figure 1.1 Genomic Structure of *PAX6*.**

Genomic organisation of the *PAX6* gene. *PAX6* is located on chromosome 11p13, and in this diagram the *PAX6* exons are shown as boxes. The diagram also shows 3 alternative promoters of *PAX6*, which are labelled P0, P1 and Pα. The major *PAX6* start codon is contained at the end of exon 4 and the stop codon is at the start of exon 13 (remainder of exon 13 coloured red to illustrate the 3' untranslated region).

### 1.1.1 Alternative promoters.

Two promoters, P0 and P1, of mouse *Pax6* have been identified (as illustrated in Figure 1.1) that direct the expression of different transcripts to different regions of the developing eye. P0 transcripts are found in the lens placode and corneal and conjunctival epithelia, whereas P1 transcripts are found in the lens placode, optic vesicle and the central nervous system (Xu *et al.*, 1999b). P0 is regulated by upstream regulatory elements, one of which directs expression of P0 transcripts to pancreatic cells (4.6kb upstream of P0, *U1* Figure 1.2), the other directing expression to ectodermally derived eye structures, namely the lens, cornea, and conjunctival epithelia (3.6kb upstream of P0, *U2* Figure 1.2) (Williams *et al.*, 1998; Kammandel *et al.*, 1999). The P1 promoter directs expression of its transcripts to the optic vesicle, central nervous system

(Xu *et al.*, 1999b) and to the developing nasal epithelium (Plaza *et al.*, 1995b); however, an upstream regulatory element, in a 5kb fragment between promoters P0 and P1 (*U3* Figure 1.2), has been shown to direct expression of P1 transcripts more specifically to regions of the brain (dorsal telencephalon and hindbrain) and the spinal cord (Kammandel *et al.*, 1999). It has also been demonstrated in quail *Pax6* that there is a temporal as well as spatial difference between the transcripts produced from promoters P0 and P1, with P1 transcripts being detected before P0 transcripts specifically during neuroretinal development (Plaza *et al.*, 1995b).

An enhancer element (*U4*, Figure 1.2) and potential promoter (Pα, Figure 1.1) was also described in quail and mouse *Pax6* that is located in intron 4 (Plaza *et al.*, 1995a; Plaza *et al.*, 1995b; Kammandel *et al.*, 1999) and directs expression of *Pax6* to the neural ectoderm derived structures, neural retina, the pigmented retina layer and the iris (Kammandel *et al.*, 1999; Xu *et al.*, 1999b). More recently it has been shown that this highly conserved enhancer element specifically mediates the formation of a distal$^{high}$ to proximal$^{low}$ gradient of PAX6 (see 1.5.1) in these regions (Baumer *et al.*, 2002).

### 1.1.2 3' Regulatory elements involved in *PAX6* expression.

It has also been proposed that 3' regulatory elements are involved in the expression of *PAX6*. Two aniridia pedigrees are associated with chromosomal rearrangements 85-185kb distal to the 3'end of *PAX6* with no mutation to the *PAX6* transcription unit. It was proposed that these disruptions were most probably due to a position effect, which disrupted the normal chromosomal environment and obstructed *PAX6* expression (Fantes *et al.*, 1995). More specifically 2 further examples of aniridia were shown where small deletions more than 11kb from the 3'end of *PAX6* deleted 3' regulatory elements (*D1* Figure 1.2) that were required for PAX6 expression (Lauderdale *et al.*, 2000).

All these chromosomal rearrangements and deletions actually occur in the final intron of another gene called *ELP4* (see Figure 1.2). However, it was reported that heterozygous loss of this gene does not reproduce the aniridia phenotype seen in the original patients and so the heterozygous loss of *ELP4* alone is not the causative factor in these patients (Kleinjan *et al.*, 2002). However, *PAX6* could have distant regulatory elements in the unrelated *ELP4* gene as this has been described previously for the alpha globin (Flint *et al.*, 2001) and human growth hormone gene clusters (Bonifer, 1999). Further to this point it has been convincingly shown that there is a complex regulatory region (termed the downstream regulatory region or DRR by the author)

about 130kb downstream of the *PAX6* poly (A) addition site (*DRR*, Figure 1.2), which contains regulatory elements that have been proven to direct reporter gene expression to distinct subsets of the overall *PAX6* expression pattern (Kleinjan *et al.*, 2001).

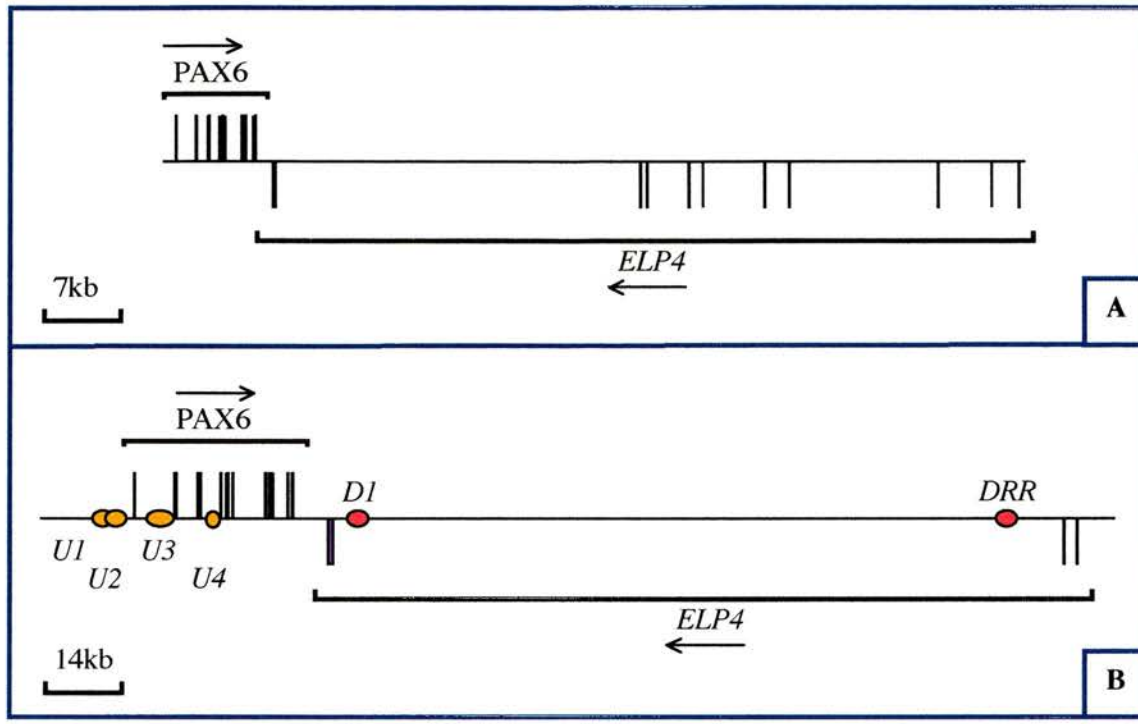



**Figure 1.2 The *PAX6* and *ELP4* loci and regulatory elements of *PAX6*.**

A. This shows the genomic organisation of the *PAX6* gene (exons 0-13, as in Figure 1.1) and the *ELP4* gene (exons 1-10 from right to left).

B. This shows an enlarged diagram of the *PAX6* and *ELP4* genes as in A, however, this time the upstream (*U1-U3*) and downstream (*D1*, *DRR*) regulatory elements have been indicated. *U1* and *U2* are 4.6kb and 3.6kb upstream of the exon 0, respectively. *U3* is situated between exons 0 and 1; *U4* is situated between exons 4 and 5. The downstream regulatory elements *D1* and *DRR* are 11kb and 130kb downstream of the 3' end of *PAX6* respectively.

The replicated function between the upstream (see 1.1.1) and downstream regulatory elements (see Figure 1.2) has led to the tentative proposal that a DNA looping model may explain the long range control of *PAX6* gene expression (Kleinjan *et al.*, 2001).

## 1.2 The Pax protein family.

The first identified paired domain containing protein was encoded by the *Drosophila* segmentation gene *paired*. Subsequently a multigene family of paired box containing genes (*Pax* genes) was identified in mammals based on sequence homology to the highly conserved paired domain of *paired* (Prosser and van Heyningen, 1998; Chi and Epstein, 2002). As already mentioned (see 1.1) human *PAX6* was first identified by positional cloning as a candidate aniridia gene (Ton *et al.*, 1991).

There are 9 members of the Pax (Paired-box) family in mammals (see Table 1.1 for a summary of PAX1-9) and these have been divided into 4 subsets due to sequence conservation, genomic structure and conserved function (Chi and Epstein, 2002). The defining feature of all Pax proteins is that they contain a 128 amino acid DNA binding domain termed the paired domain (PD), which has been conserved over evolution (Walther and Gruss, 1991). The Pax protein family is a group of transcription factors that have important functions in early development. Mutations in the genes that encode these evolutionarily conserved proteins result in acute developmental defects (see Table 1.2) in organisms ranging from flies, to mice and humans (Chi and Epstein, 2002).

### 1.2.1 Evolution and conservation of PAX6.

It was initially thought that the evolution of the eyes in vertebrates and the structurally different insect compound eye had different origins; however, with the function of PAX6 being illuminated in eye development there is now more support for the hypothesis that eyes have arisen once in evolution (Gehring and Ikeo, 1999). Heterozygous mutations in the *PAX6* gene were shown to cause similar phenotypes in mouse and humans: the Small eye phenotype in mice (Hill *et al.*, 1991) and aniridia in humans (see 1.7 for more detail). The murine and human PAX6 proteins are virtually identical in amino acid sequence and only differ by one amino acid in the alternative splice form +5a (Glaser *et al.*, 1992) (see 1.3.2). A *PAX6* homolog was also identified in *Drosophila* that transpired to be the previously described gene *eyeless*. This gene showed high sequence homology to the mouse and human proteins in the paired and homeodomains and mutation in the gene again lead to eye defects (Quiring *et al.*, 1994). This finding that *PAX6* is required for the formation of the eye in both vertebrates and invertebrates has led to the proposal that it is a master control gene that initiates eye development (Quiring *et al.*, 1994).

### 1.2.2 Ectopic expression studies.

The proposal of a monophyletic origin of the eye, notwithstanding the considerable divergence in eye morphology, has been backed up even further by ectopic expression studies. The targeted expression of the *eyeless* gene in various imaginal discs of *Drosophila* have shown that this is sufficient to produce ectopic eye structures on wings, legs and antennae (Halder *et al.*, 1995). Furthermore, it has been shown that squid (*Loligo opalescens*) Pax6 which has significant differences in sequence outside the paired domain and homeodomain, is also sufficient to produce ectopic eye structures when expressed in *Drosophila* imaginal discs (Tomarev *et al.*, 1997). The ascidian *Phallusia mammillata* is an organism that is characteristic of the ancestor from which vertebrates may have arisen and it contains a single photosensitive ocellus, which corresponds to a primeval eye structure. It is interesting to note that this urochordate PAX6 is also capable of inducing ectopic eye structures when expressed in *Drosophila* imaginal discs (Glardon *et al.*, 1997). It has also been shown that Pax6 is able to direct the formation of small but fully differentiated ectopic eyes in the vertebrate *Xenopus laevis* (Chow *et al.*, 1999). The combined information of ectopic expression studies has cemented the evolutionary conserved role of PAX6 to direct eye morphogenesis.

### 1.2.3 PAX6 DNA binding motifs and their distribution in the genome.

A consensus DNA binding sequence that is bound by the paired domain of all Pax proteins was originally revealed from detailed studies of Pax5 (Czerny and Busslinger, 1995). It was further shown that Pax6 binds to a subset of Pax5 DNA binding sites (Czerny and Busslinger, 1995). The Pax6 paired domain characteristically binds to a 16-20 base pair sequence ((ANN)TTCACGC$^{A}/_{T}$T$^{C}/_{G}$ANT$^{G}/_{T}{}^{A}/_{C}$N($^{T}/_{C}$)) whereas the homeodomain has been shown to recognise a short DNA sequence (TAAT) (Zhou *et al.*, 2002). Zhou *et al.* screened single-copy human genomic DNA libraries using cyclic amplification of protein binding sequences and electrophoretic mobility shift assays (Zhou *et al.*, 2000; Zhou *et al.*, 2002). These studies revealed that certain Pax6 binding sites occurred in sequences of the repetitive rodent B1 elements or the primate Alu elements; however, the functional significance of this observation is unclear. It has been shown that the paired domain and the homedomain can both function independently or together to regulate different genes (Zhou *et al.*, 2002). It has also been shown that the +5a PAX6 isoform has a novel binding site on DNA that is not bound by the –5a PAX6 isoform. This novel +5a PAX6 binding site contains 2 copies of an imperfect 11 base pair repeat: ATGCTCAGTGA | ATGTTCATTGA (Epstein *et al.*, 1994). It is difficult therefore to illustrate a genome wide distribution of PAX6 DNA binding motifs as although a consensus sequence has

been described, other DNA binding sites have been identified that do not share any similarity to this paired domain consensus binding site (Zhou *et al.*, 2002).

| Grp | Gene | Structure | Description |
|---|---|---|---|
| 1 | *PAX1* | PD - | Thought that these 2 members have both distinct and overlapping functions in the development of the skeleton. |
| | *PAX9* | PD - | |
| 2 | *PAX2* | PD 1 H | *Pax2* is expressed in the developing kidney, optic stalk and regions of the brain and CNS. In kidney Pax2 induces *WT1* expression, however, *Pax2* expression is repressed by WT1. |
| | *PAX5* | PD 1 H | Involved in B cell development and regulating transcription of B cell specific genes. Interaction of Sox4 and Pax5 possibly regulates B cell specific gene expression. |
| | *PAX8* | PD 1 H | Involved in directly regulating thyroid specific genes |
| 3 | *PAX4* | PD HD | Involved in specifying pancreatic cell types and regulating transcription of pancreatic hormone genes. |
| | *PAX6* | PD HD | Main function in developing eye but also involved in specifying glucagon-producing α cells in the pancreas. Interaction of Sox2 and Pax6 regulates crystallin in the developing lens. |
| 4 | *PAX3* | PD HD | Involved in formation of the heart as well as intestinal ganglia formation and development of skeletal muscle. Interaction of Sox10 and Pax3 regulates c-RET and MITF in the neural crest. |
| | *PAX7* | PD HD | Thought to specify muscle satellite cells that can be triggered to differentiate into new muscle. |

**Table 1.1 The *Pax* gene family.**

Overview of the 9 members of the *Pax* gene family, arranging them into their 4 subgroups. All contain a paired domain (PD), and 4 contain a homeodomain (HD). 3 proteins contain only the PD and the first helix of the HD (1 H). There is also a brief description of the important functions of each member of the Pax family but these are by no means exhaustive (Chi and Epstein, 2002).

| | Disease Phenotype in mice | Disease Phenotype in human |
|---|---|---|
| *Pax1* | *Pax1* mutations cause the undulated mutation that show a phenotype similar to a severe form of spina bifida (Balling *et al.*, 1988). | Mutation in the *PAX1* paired domain found in a patient with spina bifida, a neural tube defect caused as the spine doesn't close during development (Hol *et al.*, 1996). |
| *Pax2* | Human phenotype similar to Krd (Kidney, renal defects) mutant mice, that have mutations in *Pax2* and other genes (Keller *et al.*, 1994). | Mutations in *PAX2* cause kidneys to develop abnormally. Patients also exhibit an eye phenotype such as optic nerve or retinal dysplasia (Devriendt *et al.*, 1998). |
| *Pax3* | *Pax3* mutations cause the splotch mutation that show white spotting on the belly and back and have nervous system defects (Mancino *et al.*, 1992). | Mutations in *PAX3* cause Waardenburg's Syndrome I/III that produce fused digits, limb abnormalities, anophthalmia and hearing and ear defects (Hoth *et al.*, 1993). Also PAX3 FKHR fusion, see *PAX7*. |
| *Pax5* | - | A translocation alters *PAX5* transcription that may contribute to large cell lymphoma (Busslinger *et al.*, 1996). |
| *Pax6* | *Pax6* mutations cause the small eye mutation in mice that have various eye, nose and brain defects (Hill *et al.*, 1991). | Aniridia is the classical disease associated with *PAX6* mutations, see section 1.7 for more information. |
| *Pax7* | - | A translocation causing the fusion of *PAX7* to *FKHR* (forkhead), produces the paediatric soft tissue cancer alveolar rhabdomyosarcoma, a fast growing highly malignant tumour, which affects the muscles of the extremities or trunk (Whang-Peng *et al.*, 1992). |
| *Pax8* | - | Mutations in the paired domain of *PAX8* cause thyroid dysgenesis in hypothyroidism (Macchia *et al.*, 1998). |
| *Pax9* | *Pax9* loss in caudal portion of Danforth's Short tail mutant cause fewer caudal vertebrae (Neubuser *et al.*, 1995). | Mutations in PAX9 cause oligodontia, in which 6 or more of the teeth are missing (Stockton *et al.*, 2000). |

**Table 1.2 *Pax* mutations and related disease phenotypes in mice and men.**

This table illustrates the disease phenotypes that have been linked to mutations in the *Pax* genes in both mice and men. No disease phenotypes linked to *Pax4* mutations, therefore, not included.

## 1.3 The PAX6 protein.

The transcriptional start site for the PAX6 protein is located at the 3' end of exon 4. The PAX6 protein is coded from this start site in exon 4 up to the stop codon in exon 13 and there is an alternative spliceoform (5a) between exons 5 and 6. The arrangement of the PAX6 protein domains as compared to the relevant exon boundaries of the *PAX6* gene is illustrated in Figure 1.3.

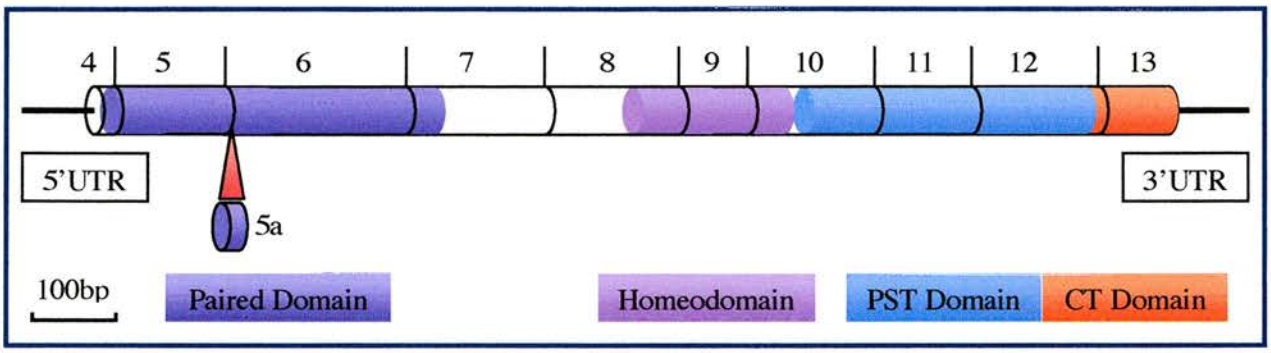



**Figure 1.3 Structure of the *PAX6* cDNA.**

This illustrates the *PAX6* cDNA with the exon boundaries marked. PAX6 has 2 main isoforms and the +5a isoform has an additional 14 amino acids that is also shown (Glaser *et al.*, 1992). The domain structure of PAX6 is also indicated, illustrating the paired domain, homeodomain, PST domain and CT (C terminal) domain as differently coloured regions of the cDNA map.

### 1.3.1 Composition and domains of PAX6.

The two major isoforms of PAX6, -5a and +5a (see 1.3.2), are 422 and 436 amino acids proteins that are composed of 3 main domains including a paired domain at the N terminus, a homeodomain and a proline, serine and threonine (PST) rich region (Ton *et al.*, 1991; Glaser *et al.*, 1992) (see Figure 1.4). The paired domain (128 amino acid domain) is a highly conserved DNA binding region that is composed of 2 subdomains, each of which contains 3 α-helices in a helix-turn-helix motif (Xu *et al.*, 1995; Xu *et al.*, 1999a). Both of the subdomains of the paired domain bind to individual consensus DNA sequences. The homeodomain (61 amino acid domain) is another DNA binding region and is also composed of 3 α-helices in a helix-turn-helix motif (Wilson *et al.*, 1995). These two DNA binding domains are highly conserved through evolution, illustrated by only a single amino acid change in both domains between PAX6 of human and zebrafish (Glaser *et al.*, 1992). The PST rich region (final 145 amino acid region) acts as a transcriptional activator (Glaser *et al.*, 1994), which together with the occurrence of the

2 specific DNA binding domains, lends support to the idea that PAX6 is a transcription factor. There is now a good deal of functional evidence to support this (see 1.6.2).

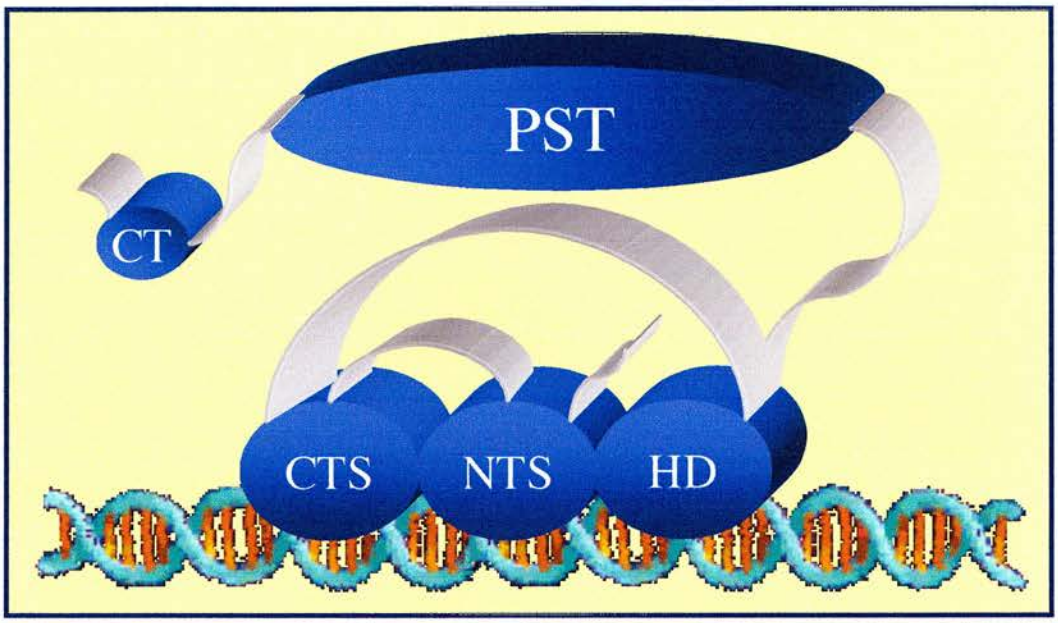



**Figure 1.4 Diagrammatic representation of the PAX6 protein.**

This cartoon illustrates how PAX6 is bound to DNA by the DNA binding regions of the protein; NTS and CTS refer to the 2 subdomains of the paired domain and HD refers to the homeodomain. The PST (proline, serine, threonine rich region) region is also illustrated and the very C terminal region (CT) is shown as a separate domain.

### 1.3.2 Isoforms – *PAX6* +/- 5a.

The PAX6 +5a isoform has an additional 14 amino acids inserted into the N terminal subdomain (NTS) of the paired domain. These added amino acids are inserted into the most conserved region of the paired domain (Walther and Gruss, 1991). It has been reported that the addition of these extra amino acids is sufficient to alter the binding characteristics of the paired domain by allowing the DNA recognition to occur through the C terminal subdomain (CTS) of the paired domain rather than the NTS (Epstein *et al.*, 1994). This has the knock-on effect that the transactivation ability through the CTS of the PAX6 +5a isoform is enhanced. This finding has led to the hypothesis that exon 5a might act as a switch that selects and specifies target genes (Azuma *et al.*, 1999). Not only do the +/- 5a isoforms of PAX6 have specific binding characteristics and specific functions during development, but also the specific ratio of +/- 5a *PAX6* transcripts is very important in eye development. This is highlighted by 2 different mutations which are sufficient to cause unusual eye phenotypes: a mutation in the exon 5a splice

acceptor site, which increases the +5a isoform and decreases the –5a isoform (Epstein *et al.*, 1994); and a missense mutation (V54D) in the alternative +5a exon which has the reciprocal effect (Azuma *et al.*, 1999).

The specific role of the PAX6 +5a isoform is still unclear; however the fact that this splice variant is only found in vertebrates could mean that it may contribute to the formation of a more complex eye structure. Exon 5a null mice have been generated and these develop with iris hypoplasia and defects of the cornea, lens and retina (Singh *et al.*, 2002). Conversely, transgenic mice that over express the PAX6 +5a isoform have also been shown to develop eye anomalies (Duncan *et al.*, 2000), increasing support for the necessity of a tightly controlled ratio of +/- 5a PAX6 isoforms in the developing eye. The requirement of the +5a isoform in iris development has led to the hypothesis that perhaps it was the evolution of this isoform that has led to more advanced features of the vertebrate eye, such as the contractile iris (Singh *et al.*, 2002). However, the +5a isoform is only expressed at low levels in the developing eye but its expression is up-regulated after birth and so this suggests that the +5a isoform may play a role in the differentiation and maintenance of the adult eye structures (Singh *et al.*, 2002).

### 1.3.3 –PD *PAX6* isoform.

It has been shown in 3 different reports that another PAX6 protein isoform (termed –PD PAX6) can be formed that is lacking the paired domain. It was shown that the *Caenorhabditis elegans Pax6* locus was capable of encoding a protein which is lacking the paired domain (Chisholm and Horvitz, 1995). It was identified that this –PD *Pax6* was expressed from a promoter in the intron before exon 8, between the paired domain and homeodomain (Zhang and Emmons, 1995). The –PD Pax6 protein was shown to be expressed in *C. elegans* with antibodies designed against the homeodomain, and this expression was both nuclear and cytoplasmic (Zhang *et al.*, 1998).

Two identical –PD Pax6 isoforms have also been identified in mouse eye and brain RNA by a modified RT-PCR technique (Gorlov and Saunders, 2002). These isoforms are missing exons 2-7 and are proposed to be generated from the 2 different *Pax6* promoters P0 and P1 (see Figure 1.1), which generate differently spliced isoforms but ultimately encode identical 221 amino acids proteins that are lacking the paired domain of Pax6 (Gorlov and Saunders, 2002).

Isoforms identified by RT-PCR must be treated with some suspicion as it still remains to be shown whether these isoforms will ever be translated into protein *in vivo*. However, another

study investigating protein isoforms of PAX6, but this time in quail neuroretina cells, has revealed a similar isoform lacking the paired domain (Carriere *et al.*, 1993). This –PD Pax6 isoform has been shown to be expressed in avian neuroretina cells by antibodies raised against the homeodomain and PST domains of Pax6 (Carriere *et al.*, 1993). Unlike the better characterised isoforms of PAX6 (+/- 5a) which are localised solely in the nucleus, it has been shown the –PD PAX6 isoform is nuclear (due to a conserved nuclear localisation signal in exon 8) but with a significant cytoplasmic localisation (Carriere *et al.*, 1995).

These combined results certainly suggest that there may well be a PAX6 isoform that is lacking the N terminal 201 amino acids that encompass the paired domain and that this isoform is also partially localised in the cytoplasm.

### 1.3.4 Two speculative *PAX6* splice variants.

Other splice variants have been reported, however, the validity of these is uncertain. One such isoform is meant to produce *PAX6* that contains exon 5a, but using an alternative 5' splice site for exon 6, whereby only 15bp of exon 6 would be included in this isoform (Gronskov *et al.*, 1999). Although this isoform has been identified from human cerebellum and lymphoblastoid cells from patients, there is still doubt about its functional significance as it has only been identified by RT-PCR. The second isoform was isolated by the same authors (Gronskov *et al.*, 2001) and produces *PAX6* that skips exon 12 which causes a shift in the read frame of the transcript and so if translated would produce a protein that was 55 amino acids longer than the 'normal' PAX6 isoform. Once again it has been shown by RT-PCR that this transcript is produced in human foetal brain cells and lymphoblastoid cell lines. Therefore, both these isoforms may be present as RNA in very small quantities in cells; however they may never be translated into protein and never contribute to the functioning of PAX6.

## 1.4 Development of structures that express *PAX6*.

This section summarises the development of the eye and the brain so that in section 1.5, which describes the expression pattern of *PAX6* in these structures, it will be possible to clearly visualise the expression pattern.

### 1.4.1 Development of the eye.

Eye development begins at the end of gastrulation (the process where the germ layers are formed and the basic body plan is established) when the forebrain neural fold forms a pair of shallow grooves called the optic sulci. Each optic sulcus then enlarges to form optic grooves and as the neural folds fuse these grooves are pushed outwards (see Figure 1.5). The optic sulci grow towards the surface ectoderm to form optic vesicles while the proximal end of the optic vesicle narrows to form the optic stalk, which connects to the brain cavity.

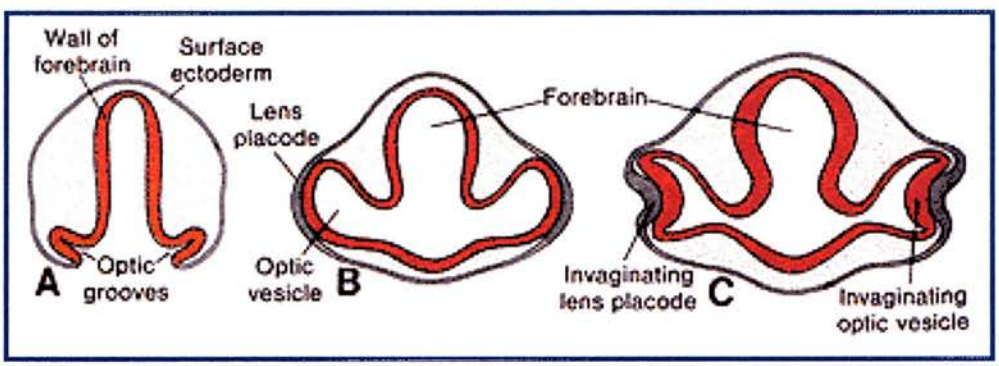



**Figure 1.5 Various developmental sections through a human forebrain.**

This figure shows transverse sections through a human forebrain at 3 stages of developments: 22 days (A), 28 days (B) and 32days. *Figure taken from the website* www.vision.ca/eye/lobby.html

Thickening of the surface ectoderm in the region apposed to the extended optic vesicle forms the lens placode. Both the lens placode and the optic vesicle invaginate at this stage and form the lens vesicle and optic cup respectively (see Figure 1.6). The lens vesicle contained within the optic cup will eventually become the solid lens of the eye. The inner and outer layers of the optic cup give rise to the thicker neural retina and thin pigmented retina respectively (Cvekl and Piatigorsky, 1996).

The invagination that forms the lens vesicle and optic cup also occurs on the surface of the optic stalk forming a linear groove called the choroidal fissure. This choroidal fissure allows the

developing lens and retina to be supplied with blood by the hyaloid artery. Ultimately the edges of the choroidal fissure will fuse, enclosing the hyaloid artery and vein as well as axons from the neural retina, which will convert the optic stalk into the optic nerve.

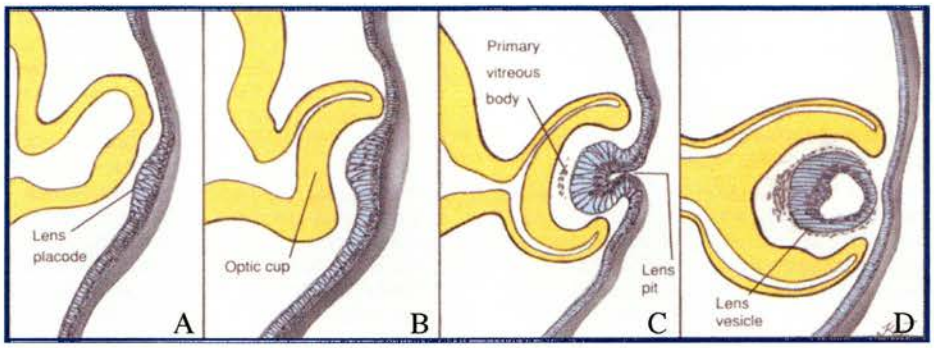



**Figure 1.6 Formation of lens placode and lens vesicle in a human eye.**

This figure illustrates the thickening of the surface ectoderm to form the lens placode (A, B) and its subsequent invagination to initially form the lens pit and then the lens vesicle (C, D). This figure also illustrates the invagination of the optic vesicle to form the optic cup (B to D). Each panel (A to D) represents a consecutive developmental stage (30 to 32 days) (Larsen, 1993).

Following the formation of the main eye structures a layer of mesenchyme surrounds the optic cup. This mesenchyme in turn will form the two layers that surround the eye, the thin inner choroid and the fibrous outer sclera. These mesenchymal layers also form over the developing lens; however, at this region the 2 layers are separated by a space called the anterior chamber. The external wall of this chamber along with surface ectoderm will form the cornea, whereas the inner wall overlying the lens will eventually develop into the pupil and vacuolisation will form the posterior chamber. Finally, the rim of the optic cup and the overlying choroid will develop into the iris and the ciliary body (Larsen, 1993). The anatomy of the mature eye is illustrated in Figure 1.7.

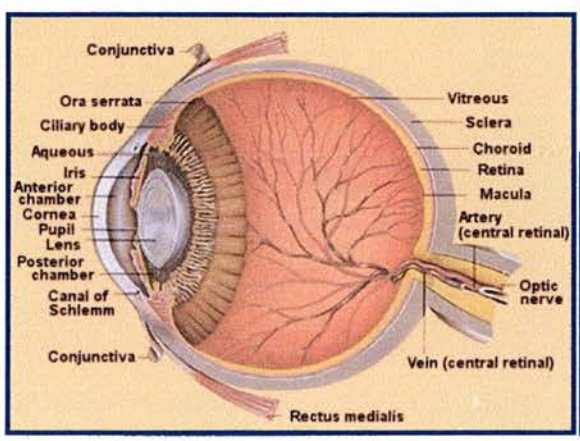



**Figure 1.7 Normal Eye.**

This diagram illustrates a vertical section through a normal eye, with the key structural features labelled. *Figure taken from the website* www.discoveryfund.org

### 1.4.2 Development of the brain.

The 3 main divisions of the developing human brain, forebrain (prosencephalon), midbrain (mesencephalon) and hindbrain (rhombencephalon) are apparent as thickened regions of the neural tube even before the start of neurulation (4 weeks gestation). Neurulation is the name given to the conversion of the neural plate into the neural tube. First a crease (neural groove) appears down the centre of the neural plate that is used to hinge the folding of the thick neural folds (see Figure 1.8).

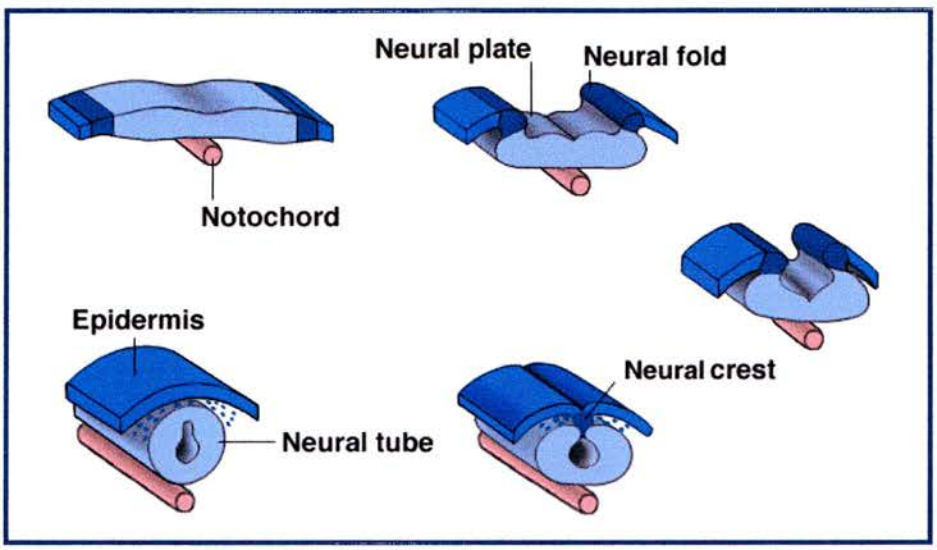



**Figure 1.8 Neurulation.**

Ectoderm in the dorsal midline thickens forming the neural plate. The edges of the neural plate roll toward each other to form the neural folds. The neural folds then meet dorsally forming the neural tube & overlying ectoderm. The neural tube will develop into the central nervous system (CNS, light blue) and the neural crest will develop into the peripheral nervous system (PNS, dark blue). *Figure taken from* www.lsic.ucla.edu:8027/.

The edges of the neural folds then fuse, first near the middle of the embryo and then proceeding rostrally (towards head) and caudally (towards tail), to form the neural tube (see Figure 1.8 and 1.9). The rostral portion of the neural tube folds at 3 points (labeled a-c, see Figure 1.9) whilst continuing to expand.

During the fifth week of brain development the forebrain is further subdivided into the telencephalon and diencephalon and the hindbrain is divided into the metencephalon and myelencephalon (see Table 1.3).

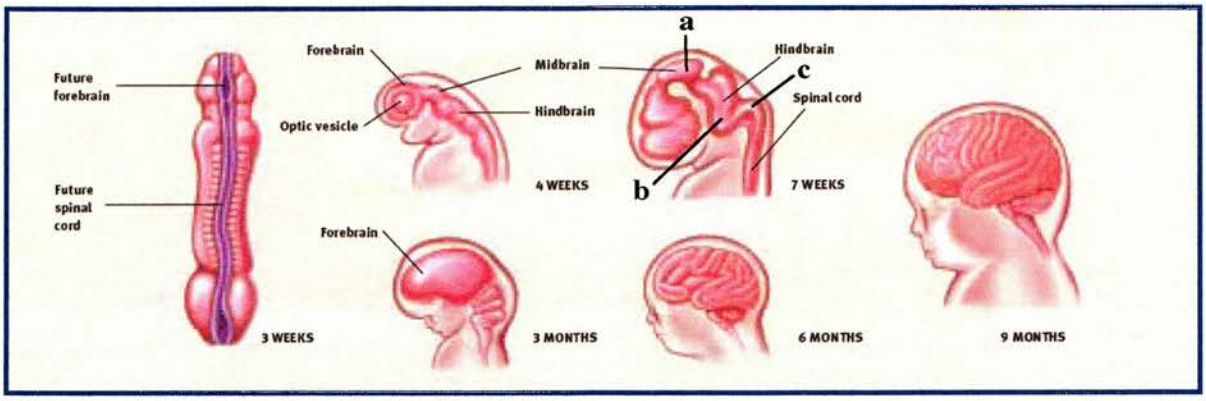



**Figure 1.9 Summary of human brain development.**

The neural tube begins to close at 3 weeks of gestation. By week 4 the presumptive major regions of the brain can be recognised, including the forebrain, midbrain, hindbrain, and optic vesicle (see section 1.4.1 for eye development). Flexing of the brain occurs at 3 points (labeled a-c in 7 week panel). Expansion and flexing of the brain then continues, with ridges of the adult brain apparent by six months (Ariniello, 2002).

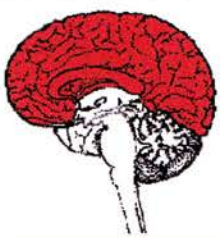
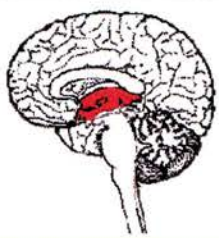
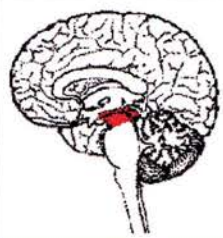
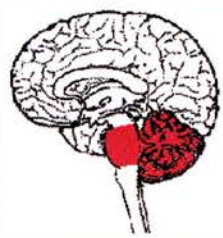
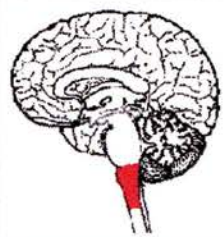
| Forebrain *(Prosencephalon)* | | Midbrain *(Mesencephalon)* | Hindbrain *(Rhombencephalon)* | |
|---|---|---|---|---|
|  |  |  |  |  |
| **Telencephalon** | **Diencephalon** | **Mesencephalon** | **Metencephalon** | **Myelencephalon** |
| Develops into: | Develops into: | Develops into: | Develops into: | Develops into: |
| Cerebral Cortex | Thalamus | Tectum | Pons | Medulla |
| Basal Ganglia | Hypothalamus | Tegmentum | Cerebellum | oblongata |
| Hippocampus | Pineal gland | Superior and | | |
| Amygdala | Pituitary | inferior colliculi | | |

**Table 1.3 Illustration of regions of the adult human brain.**

This table illustrates the regions of the adult brain that develop from different embryonic brain structures (highlighted in red). *Pictures taken from* http://faculty.washington.edu/chudler/nsdivide.html.

In brain development neurons are generated in an inside out manner to produce a highly organised structure of 6 neuronal layers. Each layer is composed of specialised neurons that have specific phenotypes and synaptic connections (Rakic and Lombroso, 1998). The neurons are originally produced from dividing progenitors in the ventricular zone and then migrate along the fibers of radial glial cells to settle and differentiate in the cortical plate (see Figure 1.10). Each subsequent set of migrating neurons will follow the same route by way of the previously formed neurons. Hence the inside out pattern of development in which the deepest cortical layers are populated by the neurons that were formed first.

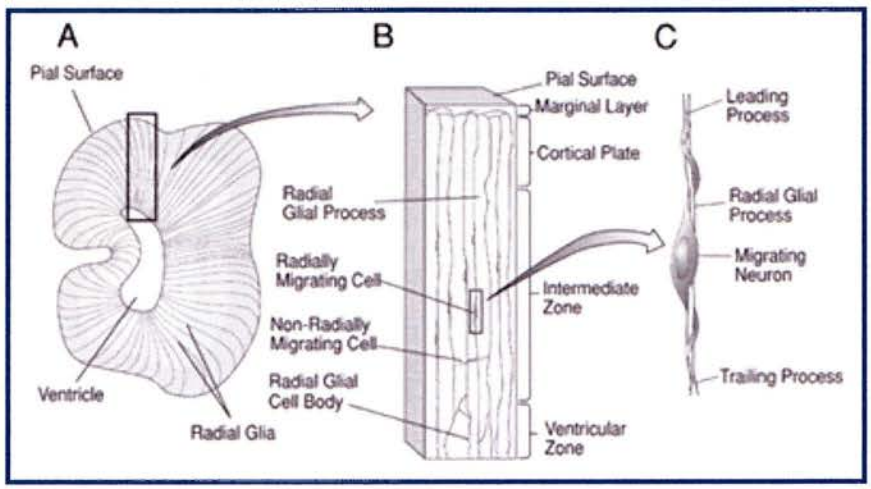



**Figure 1.10 Development of the cerebral cortex.**

A, This shows a cross-section of the occipital lobe of a developing primate brain, showing the pattern of radial glial cells that extend from the inner (ventricle) to outer (pial) surface of the brain. B, This magnified region illustrates the migration of the neurons from their birth place in the ventricular zone through the intermediate (or mantle) zone to their destination at the interface between the marginal zone and the developing cortical plate. C, This shows a single migrating neuron using a glial fibre as a guiding scaffold (Rakic and Lombroso, 1998).

After the neurons are correctly positioned they form distinct axonal connections with related regions of the cortex. The generation of specific brain regions has been suggested to be caused by the expression of certain genes and this is beginning to be clarified, for example the reciprocal gradient of *Emx2* and *Pax6* expression in the developing brain (see Figure 1.14) is thought to be involved in brain arealisation (Bishop *et al.*, 2000). The 4 main regions of the human brain and more specific regions of the forebrain, midbrain and hindbrain can be seen illustrated in Figure 1.11.

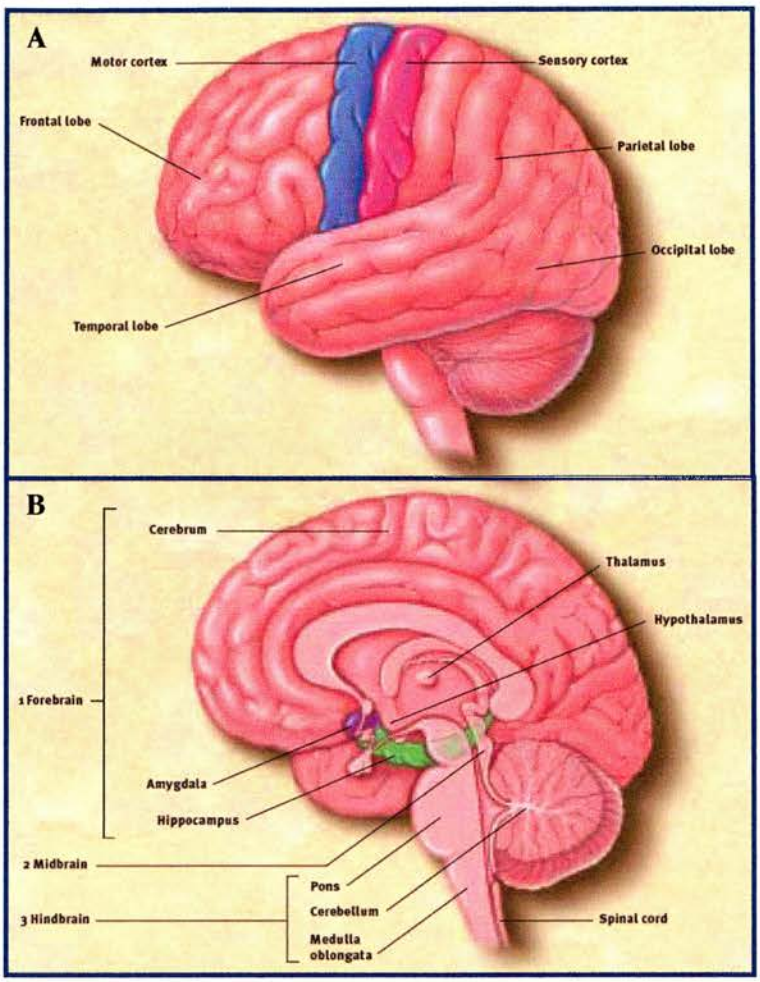



**Figure 1.11 Regions of the human brain.**

This illustration shows the important regions of a human brain. A, This shows the 4 main sections of the cerebral cortex: the occipital lobe, the temporal lobe, the parietal lobe and the frontal lobe. B, This shows a cross-section of the brain, illustrating the major internal structures in the forebrain (1), midbrain (2) and hindbrain (3) (Ariniello, 2002).

## 1.5 Where is PAX6 expressed?

The *PAX6* gene is expressed early in development, mainly in the developing eye but also in regions of the brain, central nervous system (CNS), nasal epithelium and pancreas (Prosser and van Heyningen, 1998). Although it has been shown that *Pax6* is expressed in the developing olfactory bulb and in a small subset of cells in the developing pancreas, I will concentrate on the *Pax6* expression pattern in the developing eye and brain, which is more pertinent to my project.

### 1.5.1 In the developing eye.

*PAX6* is expressed in all the major structures of the developing eye. The majority of the developmental studies that investigate the expression of the *Pax6* gene have used mouse embryos as a model system as they represent a more readily available resource than human samples. However, the expression pattern of *PAX6* that has been studied in human tissues has been shown to be interchangeable with the mouse *Pax6* expression pattern.

In the developing mouse eye, *Pax6* is first detected at day 8 of gestation in the optic sulcus and then in the resultant optic vesicle and optic stalk (Walther and Gruss, 1991). The overlying surface ectoderm also expresses *Pax6*, particularly in the thickening region, the lens placode. Between days 10-15 of gestation *Pax6* is also expressed in the optic cup, more specifically in the inner layer of the optic cup (the developing neural retina), the lens and the cornea (Walther and Gruss, 1991). In the newly formed optic cup (at 10.5 days of gestation in mice) a distal$^{high}$ to proximal$^{low}$ gradient of Pax6 expression is produced in the developing neuroretina, which is specifically mediated by the Pax6 activated, conserved intronic enhancer α (Baumer *et al.*, 2002). It has been proposed that the peak of this α initiated Pax6 expression, at the anterior rim of the optic cup, corresponds to the future iris epithelium (Baumer *et al.*, 2002).

At early stages of development *PAX6* is expressed in the entire developing retina; however, at later stages of development this expression pattern is restricted to the ganglion cell layer and the inner and outer portions of the inner nuclear layer, which will differentiate into amacrine and horizontal cells (Ton *et al.*, 1991). A picture illustrating the expression of *PAX6* at this stage of human eye development is shown in Figure 1.12.

One must be careful when drawing conclusions from the RNA expression pattern of a gene, as it does not always follow that the expression pattern of the RNA will be mimicked by the

expression pattern of the translated protein. However, in the case of PAX6 it has been shown that the protein expression mimics the RNA expression in the developing eye (Nishina *et al.*, 1999). The early expression of PAX6 in human eye sections has been shown by anti-PAX6 antibody staining (Nishina *et al.*, 1999) to resemble the in situ hybridisation findings of *PAX6* expression in mice (Walther and Gruss, 1991) and a 49 day old human eye (see Figure 1.12).

**Figure 1.12 *PAX6* RNA expression in the human eye.**

This picture shows the *PAX6* RNA expression (as determined by in situ hybridisation) in a human eye at 7 weeks of gestation. *PAX6* expression can be seen in the optic vesicle and the overlying ectoderm as well as in the pigmented and neuronal retina (Ton *et al.*, 1991).

### 1.5.2 In the developing brain.

*PAX6* has also been shown to be expressed in the brain. However, this expression pattern is extremely complex. *Pax6* expression, as in the eye, begins at day 8 of gestation in the developing mouse in the neuroepithelium of the forebrain and hindbrain, and is then followed by expression along the length of the neural tube (Glardon *et al.*, 1997).

During brain development (days 10.5-18.5 of gestation) expression of *Pax6* extends from the telencephalon into the diencephalon (the 2 main regions of the forebrain) but never penetrates the boundary of midbrain (Walther and Gruss, 1991). A high level of *Pax6* expression persists in the developing ventral thalamus (Stoykova *et al.*, 1996) from 10.5 to 18.5 days of gestation, however, a much lower *Pax6* expression level is seen in the developing dorsal thalamus which is undetectable at 15.5 days of gestation (Walther and Gruss, 1991). This suggests that a gradient of *Pax6* is establish in this region, as in the neuroretina in the eye. Delayed *Pax6* expression, from 15.5 to 18.5 days of gestation, also occurs in the developing metencephalon (Walther and Gruss, 1991).

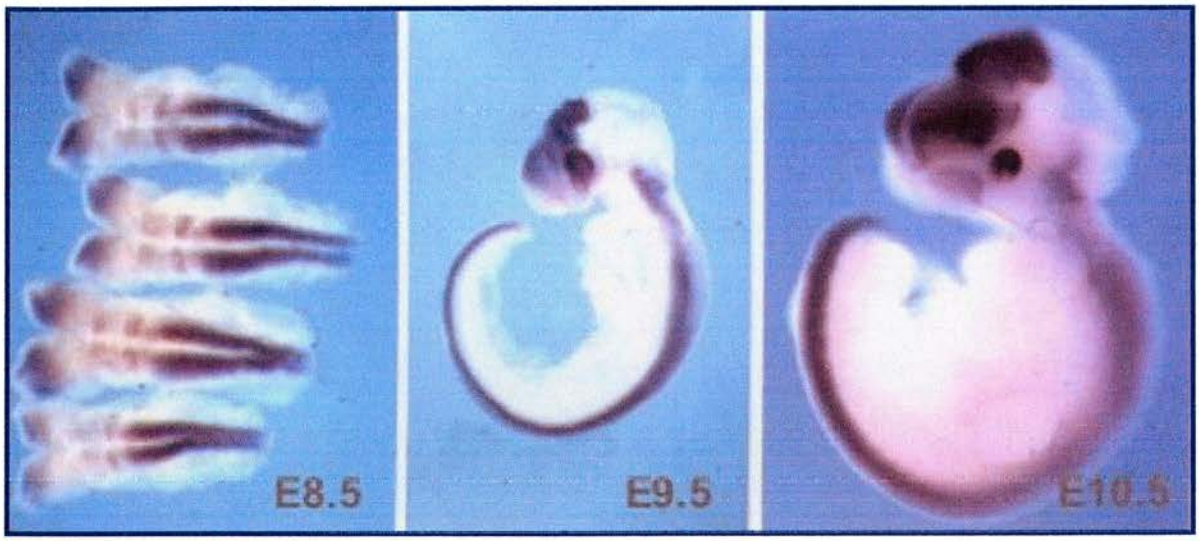



**Figure 1.13 Expression of *Pax6* in developing mouse embryo.**

*Pax6* expression can be seen in the developing forebrain and hindbrain as well as along the length of the developing neural tube. *Pax6* expression can also be seen in the developing olfactory structures and in the eye (*Picture provided by P. Rashbass*).

*Pax6* expression is detected at 13.5 days of gestation in the ventricular zone and this extends to the subventricular zone at 16.5 days of gestation (Caric *et al.*, 1997). The cells generated from these mitotically active regions then migrate from this inner region of the telencephalon to the cortical plate along radial glial cells (see Figure 1.10), which also express *Pax6* (Heins *et al.*, 2002).

*Pax6* has also been shown to be expressed in a high rostrolateral to low caudomedial gradient which is complemented by an opposing gradient of *Emx2* (Bishop *et al.*, 2002).

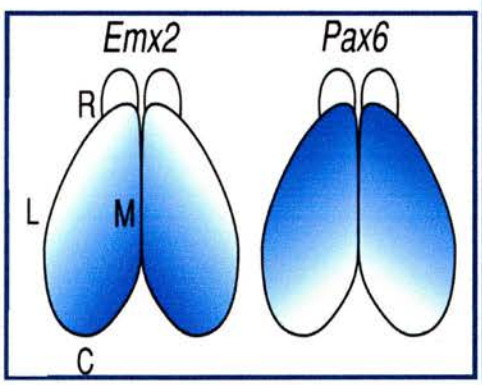



**Figure 1.14 *Pax6* and *Emx2* expression in neocortex.**

The predicted expression of *Pax6* and *Emx2* in opposing directions in the developing neocortex. It is thought that this reciprocal expression is required to define specific regions of the brain. Diagrams represent dorsal views of the mouse neocortex. C, Caudal; L, Lateral; M, Medial; R, Rostral (Bishop *et al.*, 2002).

It has been shown that the complementary function of Emx2 and Pax6 specifies area identity in the neocortex (Bishop *et al.*, 2002).

## 1.6 Function of the PAX6 protein.

I have already explained that the PAX6 protein is thought to be a transcription factor and I have outlined the development of the brain and eye, the two main structures in which PAX6 is expressed. Our knowledge of the role of PAX6 in patterning the developing brain has expanded over recent years. PAX6 is involved in regionalisation of the brain (briefly mentioned in section 1.5.2) and also in cell migration (Talamillo *et al.*, 2003) and axon guidance (Mastick *et al.*, 1997) in forebrain development.

However, PAX6 is best known for its conserved role in eye development and therefore the main area of research into the functioning of PAX6 has been in the developing eye. This section is intended to highlight some specific details of how PAX6 functions at the molecular level in the development of the eye.

### 1.6.1 Regulators of *Pax6* expression in the eye.

Pax2 has been shown to bind to the regulatory region between exons 4 and 5 of *Pax6* (*U4* in Figure 1.2) and this down regulates expression, whereas the binding of Pax6 to an upstream regulatory region in *Pax2* also down regulates expression (Schwarz *et al.*, 2000). Conversely, Pax6 has been shown to be able to stimulate its own expression by binding to the same regulatory region (Schwarz *et al.*, 2000; Baumer *et al.*, 2002). As PAX2 is specifically expressed in the optic stalk and PAX6 is expressed in the developing retina, and missense mutations of PAX6 lead to optic nerve malformation (Azuma *et al.*, 2003), then it is possible that the overlapping expression of PAX2 and PAX6 set up an optic cup/optic stalk border during development. This hypothesis is backed up by information from the *cyclops* mutant in zebrafish (Macdonald *et al.*, 1995). This mutant has fused eyes and this is due to the fact that *Pax6* expression is expanded from the developing retina into the normal expression area of *Pax2* (the presumptive optic stalk) and *Pax2* expression is almost absent. Therefore, *Pax6* expression encroaches on the normal *Pax2* expression area and the Pax2/Pax6 boundary is altered.

As has been alluded to several times already it is known that PAX6 is able to regulate its own expression during development. I have already commented that binding of Pax6 to a regulatory region between exons 4 and 5 (*U4* see Figure 1.2) specifically directs expression of further *Pax6* transcripts in a gradient in the developing neuroretina (Baumer *et al.*, 2002). However, Pax6 has also been shown to bind to the enhancer element upstream of the P0 promoter (*U2* see Figure

1.2) in combination with the transcription factors Sox2 and Sox3 (Aota *et al.*, 2003). It is postulated that the positive autoregulation of Pax6 may be adjusted by this interaction with Sox2 and Sox3.

The regulation of *PAX6* expression in lens development also seems to be under the control of the homeoproteins Meis1 and Meis2. It has been shown that these proteins bind to the specific lens enhancer element (*U2* see Figure 1.2) and increased Meis2 has been shown to increase the expression levels of PAX6 in the lens (Zhang *et al.*, 2002). Specific levels of Pax6 expression are also required in the developing lens to regulate expression of the homeobox gene *Six3*, and Six3 also binds to the lens specific enhancer of *Pax6* (*U2* see Figure 1.2) to activate *Pax6* expression (Goudreau *et al.*, 2002).

Therefore, it can be seen that not only is the regulation of *PAX6* finely controlled in the developing eye, but that this controlled expression of *PAX6* will also influence other genes and proteins (see 1.6.2 and 1.6.3).

### 1.6.2 Genes regulated by the PAX6 protein.

It has already been mentioned that the binding of PAX6 protein to specific *PAX6* regulatory elements specifically stimulates expression of the gene (Schwarz *et al.*, 2000; Baumer *et al.*, 2002; Aota *et al.*, 2003). However, the Pax6 protein has also been reported to directly activate other genes that are involved in eye development. As already mentioned Pax6 regulates activation of *Six3* in a dose dependant manner in the developing lens, so the more PAX6, the more *Six3* expression (Goudreau *et al.*, 2002). Consistent with the function of the Pax6 protein in the developing lens, it also regulates expression of different *crystallin* genes (Cvekl and Piatigorsky, 1996) and *Maf* genes (Sakai *et al.*, 2001). Pax6 also controls the activation of bHLH factors such as *Ngn2*, which are required in the developing retina to produce the correct cell types from retinal progenitor cells (Marquardt *et al.*, 2001).

It is thought that *Foxe3* expression is activated by a threshold amount of Pax6 in the developing lens (Dimanlig *et al.*, 2001). Also the *Necab* gene is thought to be a downstream target of Pax6, and be involved in retina development (Bernier *et al.*, 2001). Although the expression of both these genes has been shown to be absent in mutant mice, a direct interaction with Pax6 has not been confirmed and so these predictions are still speculative.

Pax6 not only stimulates gene expression, but it also represses gene expression as has already been described (see 1.6.1) with the example that Pax6 represses *Pax2* expression (Schwarz *et al.*, 2000). Pax6 has also been shown to repress expression of a *β-crystallin* gene by blocking an activation domain in its promoter (Duncan *et al.*, 1998). The dual roles of Pax6 as a repressor and activator further the information that Pax6 is responsible for regulating a complex spatial and temporal pattern of gene expression in the developing eye.

### 1.6.3 PAX6 protein-protein interactions.

As has just been shown *PAX6* is highly regulated; not only is the expression of *PAX6* restricted to specific areas of the developing embryo, but also this has subsequent effects on the regulation of other genes. However, it is important to remember that the specific regulation of genes by the PAX6 protein is not just accounted for by the highly controlled expression pattern of PAX6 or its specific DNA binding domains. It is becoming apparent that transcription factors do not act in isolation but are dependent on interactions with other proteins to carry out their function (Oikawa and Yamada, 2003; Westerman *et al.*, 2003). These interactions, which are themselves dependant on temporal and spatial expression, introduce more specificity into the regulatory function of PAX6. To date three protein-protein interactions of PAX6 have been described: the specific interaction of Sox2 and Pax6; the interaction of mDia and Pax6; and the interaction of Maf proteins and Pax6.

As *SOX2* has a spatially and temporarily controlled expression pattern in the developing eye and central nervous system and gene mutations have been reported in patients with anophthalmia, a severe structural eye malformation (Fantes *et al.*, 2003) it is plausible to hypothesise that the interaction of Sox2 and Pax6 is vital in certain stages of development. Specifically, it has been postulated that lens development is initiated by the binding of a Sox2/Pax6 complex onto a lens specific (*δ-crystallin*) enhancer element in chicks (Kamachi *et al.*, 2001). This example substantiates the hypothesis that the Sox2/Pax6 interaction creates a specific combination of DNA binding domains that will bind to a corresponding combination of binding sites in the enhancer region (Kamachi *et al.*, 2000).

Therefore, one could envisage a scenario by which PAX6 interacts with many different proteins to allow binding to various enhancer elements to specifically regulate its function during development.

It has been found from a yeast 2-hybrid screen using mDia as bait that Pax6 interacts with this protein (Tominaga *et al.*, 2002). It has been proposed that mDia binds across the homeodomain of Pax6 in the nucleus of cerebellar granule cells, which causes Pax6 to dissociate from DNA and translocate to the cytoplasm (Tominaga *et al.*, 2002). In this way it has been shown that the transcriptional activity of Pax6 is reduced due to the binding of mDia, and this is responsible for neurite elongation and branching in the cerebellar granule cells. Therefore, it has been suggested that mDia regulates axon pathfinding and Pax6 activated gene transcription and that this could potentially provide a mechanism of how early neuronal development is modulated due to its interaction with Pax6 (Tominaga *et al.*, 2002).

Furthermore, it has been shown that Pax6 and Maf proteins interact in pull down experiments *in vitro* and that not only does Pax6 regulate expression of the pancreatic hormone gene, *glucagon*, but also the transactivation of this gene is increased by the combined binding to its promoter of both the Pax6 and Maf proteins (Planque *et al.*, 2001).

| | | More Information | Reference |
|---|---|---|---|
| *Genes directly regulated by Pax6* | *Six3* | Regulates activation of *Six3* in a dose dependant manner in the developing lens. | Goudreau *et al.*, (2002) |
| | *Crystallin* | Regulates different *crystallin* genes in the developing lens. | Cvekl and Piatigorsky, (1996) |
| | *Maf* | Activates gene expression of *c-Maf* in lens development. | Sakai *et al.*, (2001) |
| | *Ngn2* | Controls the activation of *Ngn2* in the developing retina | Marquardt *et al.*, (2001) |
| *Indirect evidence of genes regulated by Pax6* | *Foxe3* | Pax6 speculated to activate *Foxe3* expression in the lens. | Dimanlig *et al.*, (2001) |
| | *Necab* | Pax6 proposed to regulate *Necab* expression in retinal development. | Bernier *et al.*, (2001) |
| *Known protein interactions of Pax6* | Sox2 | Sox2/Pax6 complex binds to a lens specific enhancer element. | Kamachi *et al.*, (2001) |
| | mDia | mDia binds to Pax6 in cerebellar granule cells. | Tominaga *et al.*, (2002) |
| | Maf | Maf binds to Pax6 causing increased transactivation of the *glucagon* gene. | Planque *et al.*, (2001) |

**Table 1.4 Summary of gene regulation by the PAX6 protein.**

This table summarises the genes that are regulated by the PAX6 protein (see 1.6.2), and these have been split into 2 categories depending on whether there is direct or indirect evidence for PAX6 regulation. Known protein-protein interactions are also shown (see 1.6.3).

## 1.7 Mutations of PAX6.

Classically, mutation of *PAX6* was associated with a congenital eye malformation, aniridia (Greek for absence of the iris, see Figure 1.15), that is a rare, bilateral, panophthalmic disorder, characterised by iris hypoplasia (Ton *et al.*, 1991; Jordan *et al.*, 1992). Aniridia can also be associated with corneal, lens, optic nerve, and retinal manifestations as well as foveal and optic nerve hypoplasia, which cause a reduction in visual acuity and congenital sensory nystagmus. Vision may also deteriorate as patients age due to the onset of cataracts, glaucoma, and corneal opacification *(Taken from www.emedicine.com)*.

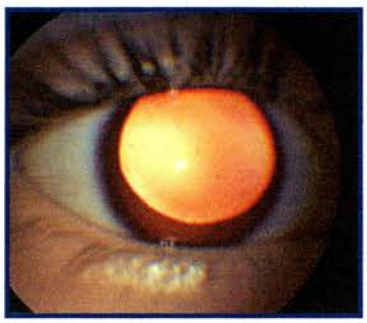


**Figure 1.15 Illustration of aniridia.**

Only a very small region of iris tissue remains at the bottom of the eye. The pupil of the eye is brightly illuminated as light is being reflected off the retina (Hanson and Van Heyningen, 1995).

Although aniridia is the most common eye disease phenotype associated with mutations in *PAX6*, other diseases such as Peters anomaly, ectopia pupillae and foveal hypoplasia are also associated with *PAX6* mutations (see Figure 1.16).

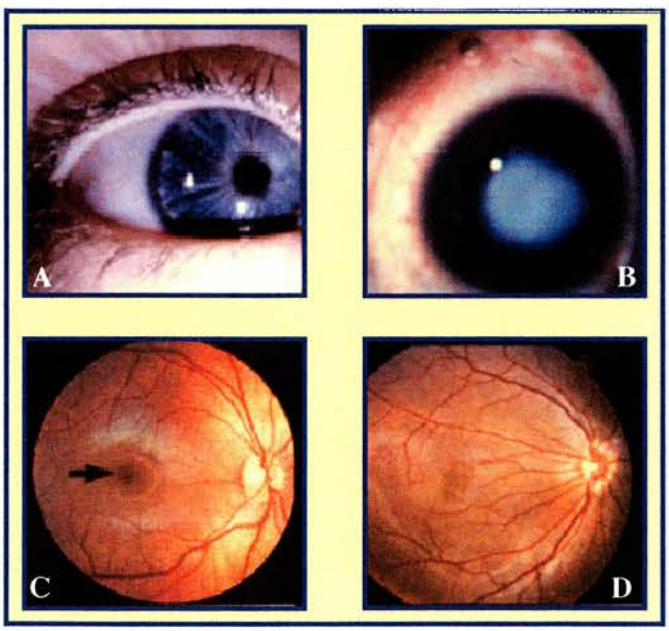



**Figure 1.16 Eye diseases associated with *PAX6* mutations.**

Illustration of eye diseases other than aniridia associated with *PAX6* mutations: A, shows ectopia pupillae, with the pupil offset towards the nose; B, illustrates Peters anomaly showing central corneal opacification; C and D illustrate foveal hypoplasia, with C, showing a normal retina and D, illustrates the disease showing a poorly formed fovea with abnormal vascularisation.

Peters anomaly is a rare form of anterior segment dysgenesis in which abnormal cleavage of the anterior chamber occurs (causing adhesions between the cornea, iris and/or the lens) and is characterised by central or paracentral corneal opacifications *(Taken from www.emedicine.com)*. Ectopia pupillae is a congenital eye malformation in which the pupils are offset from their central position, normally towards the nose (Hanson *et al.*, 1999). The fovea is the region of the retina that contains a single layer of cones with no overlapping blood vessels and has the greatest visual acuity. In foveal hypoplasia, the fovea is underdeveloped and there is vascularisation through the vestigial fovea (Azuma *et al.*, 1996).

Heterozygous mutations in *PAX6* are the genetic cause for the previously described congenital eye malformations (Prosser and van Heyningen, 1998) and it is thought that most of these mutations result in non-functional alleles (see 1.7.1). Homozygous mutations cause far more severe phenotypes. Two cases have been described: One in which the patient died at 37 weeks gestation and was completely lacking eyes, nose and adrenal gland, although it was only presumed that this phenotype was caused by a homozygous *PAX6* mutation (Hodgson and Saunders, 1980); and a patient that was a compound heterozygote that developed with no eyes, an underdeveloped brain and central nervous system abnormalities and died 8 days after birth (Glaser *et al.*, 1994).

### 1.7.1 Overview of *PAX6* mutants and the database.

As I have already mentioned heterozygous mutations of *PAX6* cause eye disease phenotypes, however just as there is a lot of variation in the disease phenotype there is also a wide range of genotypic mutations that cause these diseases. The entire spectrum of mutations that occur in the *PAX6* gene are collated in the *PAX6* Allelic Variant Database (http://pax6.hgu.mrc.ac.uk/). There are 292 mutations recorded in this database and 275 of these are mutations within the coding region of *PAX6*, which cause disease phenotypes. The variety of different types of mutation that lead to these 275 pathological mutations can be seen in Figure 1.17.

It has also been noted that certain positions in the *PAX6* gene are 'mutation hotspots' which include mutations that affect the *PAX6* stop codon and the mutations R203X (14 cases), R240X (17 cases) and R317X (12 cases) (Chao *et al.*, 2003). Although there are far more mutations of *PAX6* than are mentioned here, I am going to concentrate more on the different types of mutation, i.e. missense or null mutations, with specific reference to any interesting or relevant mutations.

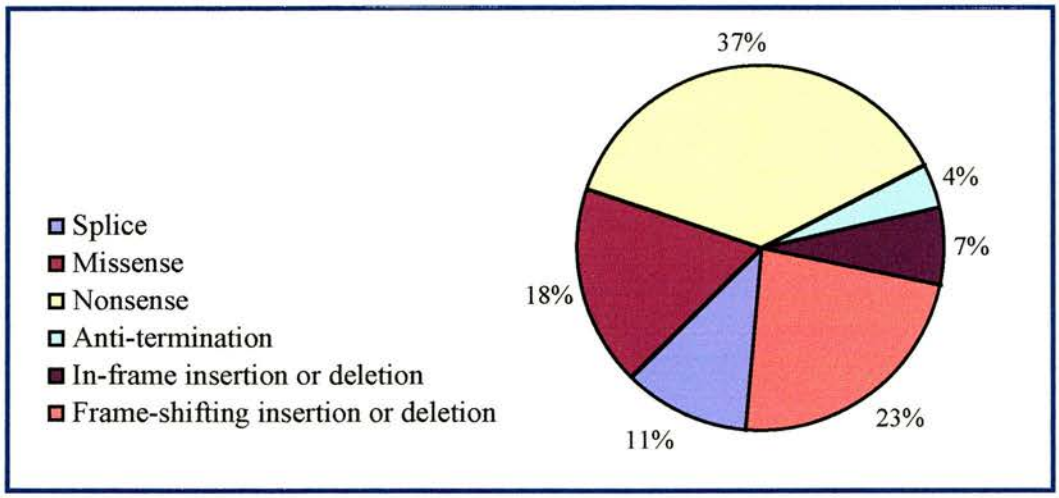



**Figure 1.17 *PAX6* mutations.**

This pie chart illustrates the frequency of the different types of pathological mutations that occur within the *PAX6* coding region. Anti-termination mutations refer specifically to mutations that change the stop codon to a coding codon (Hanson, 2003).

### 1.7.2 Null Mutations.

Classically aniridia was associated with a complete deletion of one allele of the *PAX6* gene due to the hemizygous deletion of 11p13 that typifies the WAGR syndrome (Wilms tumour, aniridia, genitourinary malformations and mental retardation) (Ton *et al.*, 1991). This led to the hypothesis that aniridia was caused by haploinsufficiency of *PAX6* and that loss of function of one allele meant that insufficient wild type protein was produced to elicit normal eye development (Martha *et al.*, 1994). It has since been shown that intragenic mutations of *PAX6* also cause eye disease phenotypes and 71% of these introduce a premature termination codon (PTC) due to either frameshift, nonsense or splice mutations. This was initially envisaged to result in the formation of truncated PAX6 proteins (see 1.7.4) although more recently it has been suggested that mutations that produce a mRNA containing a premature stop codon may never be translated into protein due to the function of nonsense-mediated decay (Vincent *et al.*, 2003).

### 1.7.3 Missense Mutations.

Missense mutations are most commonly associated with atypical aniridia or uncommon phenotypes (Vincent *et al.*, 2003) and seem to be underrepresented in the mutational spectrum of PAX6 mutations. It has been suggested that a reason for this might be that missense mutations might segregate with highly diverse or non-ocular phenotypes (Hanson *et al.*, 1999). Therefore, missense mutations may well be associated with phenotypes rather different from aniridia due to the production of a protein with altered rather than a loss of function (Axton *et al.*, 1997).

### 1.7.4 Haploinsufficiency and nonsense-mediated mRNA decay.

As discussed in section 1.7.2 haploinsufficiency has been suggested as the explanation for the heterozygous phenotype that causes aniridia, meaning that the mutated *PAX6* allele produces inactive protein and that the normal *PAX6* allele cannot make sufficient wild type protein to sustain normal eye development (Martha *et al.*, 1994). As the *PAX6* mutation database expanded it was seen that the majority (71%) of *PAX6* mutations resulted in the addition of a premature termination codon (PTC). Most of these mutants introduced a PTC before the PST domain, which would theoretically form proteins that are unlikely to have any function (Martha *et al.*, 1994) and fitting a haploinsufficient cause of aniridia. However, transcripts that added a PTC into the PST domain (see Figure 1.3 for position of the PST domain), if translated, would represent proteins that contained intact DNA binding domains devoid of the transactivating PST domain. These truncated PAX6 proteins have been shown to bind more strongly than full length PAX6 to PAX6 DNA-binding sites (Singh *et al.*, 1998) and also to show a reduced transactivation ability (Singh *et al.*, 1998; Tang *et al.*, 1998). This fuelled an alternate hypothesis that suggested that the truncated protein, devoid of the transactivating PST domain would exert a dominant negative effect by competing for PAX6 DNA binding sites and interrupt the normal functioning of PAX6 to activate downstream genes. This fitted with the information obtained by Singh *et al.*, (1998) and was backed up by further *in vivo* evidence, which showed that the formation of endogenous and ectopic eyes in *Xenopus laevis* were suppressed by the introduction of a dominant negative mutant (Chow *et al.*, 1999).

Although dominant negative effects have been reported for truncated PAX6 proteins (Singh *et al.*, 1998) it must be stressed that this hypothesis is still questionable. The highly contrived experiments that support the dominant negative effect avoid a fundamental control mechanism, nonsense-mediated mRNA decay (NMD). This is a mechanism employed by cells to selectively destroy mRNAs that contain premature stop codons, so that only full-length proteins are

produced (Maquat and Carmichael, 2001; Byers, 2002). When investigating mutations that cause premature termination codons, or frame-shift mutations that produce proteins with novel properties, proof of the mRNA stability and protein production must be demonstrated before conclusions are drawn as to the detrimental effect of the mutant protein. The experiments supporting dominant negative mutants overlooked this vital point and so the results illustrated an unnatural occurrence. This is supported by the phenotype of mutations that create premature stop codons in the PST domain *in vivo*, which seem to suggest haploinsufficiency (Vincent *et al.*, 2003) as all 44 cases of patients with PST domain mutations show aniridia or closely related eye phenotypes. If the normal function of PAX6 is almost completely blocked then severe phenotypes, more akin to homozygotes, should be seen. Therefore, it is likely that NMD prevents these proteins from being formed.

For the majority of mutations of *PAX6*, the explanation for disease phenotypes will be loss of function of one allele causing haploinsufficiency. However, certain mutations actually cause read through of the normal stop codon of *PAX6* (termed run-on mutations).

### 1.7.5 Run-on Mutations.

Run-on mutations that cause read through of the normal stop codon into the 3' untranslated region (UTR) of *PAX6* are known to cause aniridia. mRNAs that contain run-on mutations would not be expected to be degraded by nonsense-mediated decay because they don't introduce a premature stop codon. Furthermore, it has been reported that nonsense-mediated decay does not act within the final exon of a gene (Byers, 2002) so it is expected that run-on mutations will produce elongated PAX6 protein. The C terminal region of the PST domain of PAX6 has been shown to be highly conserved between species (see 1.8) and altering the C terminal amino acid can interfere with the binding capabilities of the homeodomain (Singh *et al.*, 2001). Therefore, these elongated PAX6 proteins may interfere with the normal functioning of the C terminal region and so instigate a novel function, although existing patient phenotypes associated with this type of mutation, suggest that the elongated PAX6 protein will have a partial or complete loss of function.

### 1.7.6 C terminal mutations.

21% of the *PAX6* mutations occur in the coding region of the PST domain; however, more specifically mutations do occur in a C terminal region of the PST domain of PAX6 (to be defined in more detail in section 1.8), which are capable of causing eye malformations. These mutations can be as subtle as the insertion or changing of a single nucleotide, or the deletion of 10 nucleotides from the terminal region of the PST domain of PAX6 (see Figure 1.18).

A single nucleotide change from A to G at position 1627 causes a glutamine to arginine amino acid substitution in the last codon (422) of *PAX6*. This missense mutation (Q422R, see Figure 1.18) was reported in two patients both of whom were shown to have ocular anterior segment malformations (Azuma and Yamada, 1998; Singh *et al.*, 2001). One patient developed mild aniridia characterised by uveal ectropion which is an invasion of the cornea by the conjunctival epithelium (Azuma and Yamada, 1998) and the other patient has more classical aniridia with associated foveal hypoplasia (see Figure 1.16) (Singh *et al.*, 2001).

Another specific C terminal mutation has been reported in a patient (WHEDA, see figure 1.18) suffering from aniridia in which there is a 10 base pair deletion in exon 13 of *PAX6*. This deletion occurs just before the stop codon in *PAX6*, causing read through and generating an almost normal PAX6 protein (missing last 5 amino acids) but with an elongated C terminal end that extends 103 amino acids into the 3' untranslated region (Heyman *et al.*, 1999). It is interesting to note that not only does this patient have an abnormal eye phenotype, aniridia, but they also have an unusual neurobehavioural phenotype, namely impaired social understanding and verbal inhibition, that may be a consequence of this mutation.

A C terminal mutation (X423L) has been reported in 11 patients suffering from aniridia (Baum *et al.*, 1999; Singh *et al.*, 2001; Sisodiya *et al.*, 2001; Chao *et al.*, 2003). The X423L mutation (see Figure 1.18) causes a single T nucleotide insertion into the stop codon, which changes it to a leucine and so forms a full length PAX6 protein but with an elongated C terminal end that extends for a further 35 amino acids into the 3' untranslated region. In addition to aniridia one patient, REMAR, was also missing the anterior commissure (AC), a region in the brain involved in communication between the brain hemispheres. This patient was also reported to have a severe defect in olfactory function although with an apparently normal olfactory bulb which indicates a possible defect in neuronal patterning in this region (Sisodiya *et al.*, 2001). Further evidence for defects in neuronal development due to this mutation came from the discovery that the nephew of REMAR had epilepsy and both the nephew and the sister of REMAR had a rare

brain disorder, polymicrogyria (PMG), that is thought to arise from the disruption of neuronal migration (Mitchell *et al.*, 2003). Another patient was also shown to have brain asymmetry as identified by a computer tomography scan which manifested itself as autistic-type behaviour (Chao *et al.*, 2003), which is similar to the phenotype of the 1615del10 (WHEDA) PAX6 mutation that also results in an elongated protein. However, it must be remembered that the autistic-type behaviour has only been reported in 1 patient out of 11 with the X423L mutation and so any conclusions drawn relating C terminal PAX6 elongation and autistic-type behaviour is highly speculative.

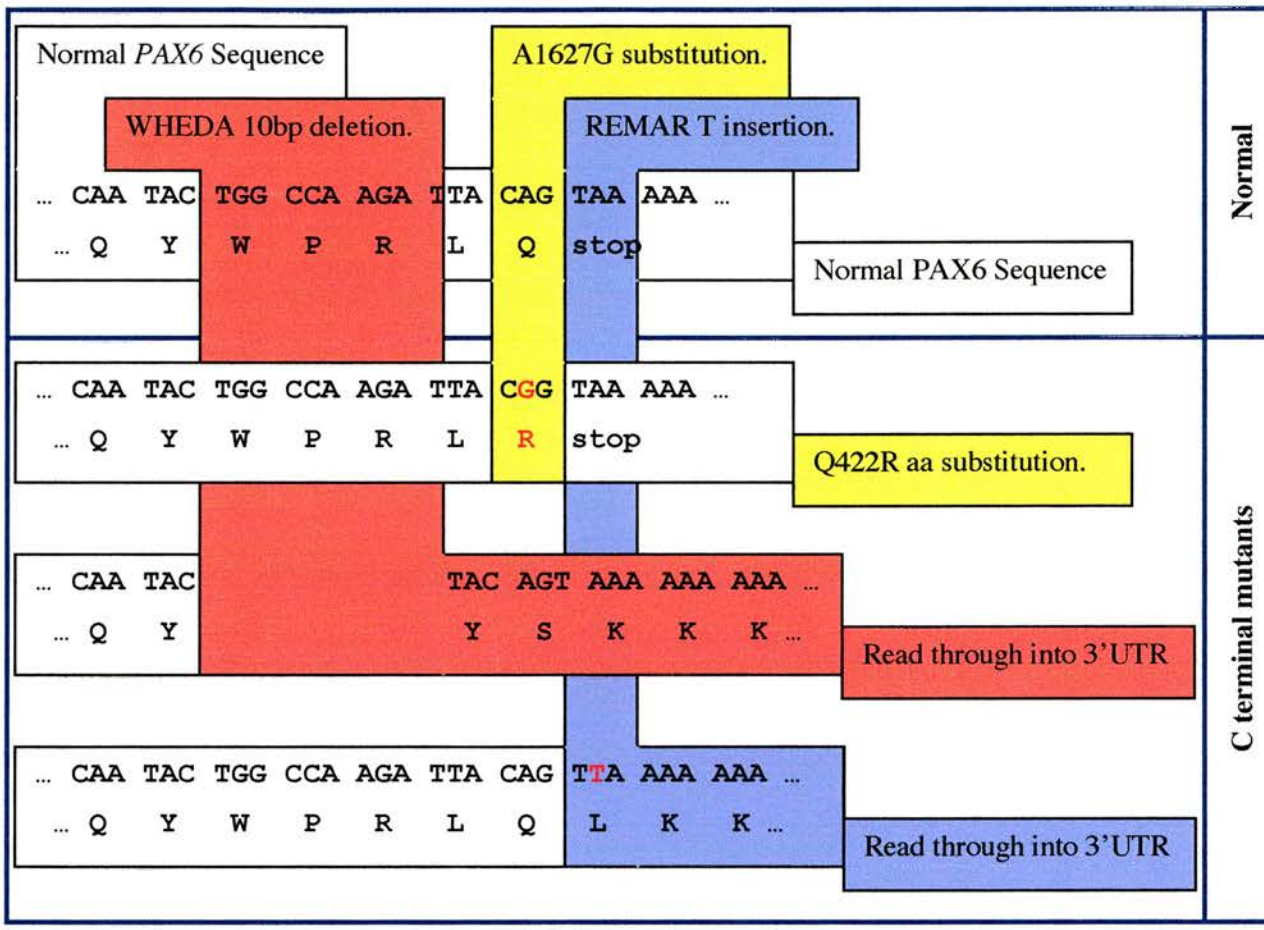



**Figure 1.18 Schematic of C terminal PAX6 mutations.**

This shows the 3 C terminal PAX6 mutants: Q422R; WHEDA (1615del10); and REMAR (X423L). The normal C terminal *PAX6* nucleotide and PAX6 amino acid sequences are illustrated at the top of this figure. The A to G nucleotide substitution and the resulting Q422R amino acid change is illustrated in yellow. The 10bp deletion of the WHEDA mutant and subsequent altered reading frame and elongated protein is illustrated in red. The substitution of the stop codon by a leucine due to the insertion of a T nucleotide into the stop codon and the subsequent elongated protein of the REMAR mutant is illustrated in blue.

# 1.8 Importance of C terminal motif.

## 1.8.1 Known homology studies.

It has been noted in a previous study that there is a region at the C terminus of the PST domain which shows very high sequence conservation between vertebrates, sea urchin (*Paracentrotus lividus*) and squid (*Loligo opalescens*) when the whole PAX6 protein sequence is used as a query in a BLAST search (Glardon *et al.*, 1997).

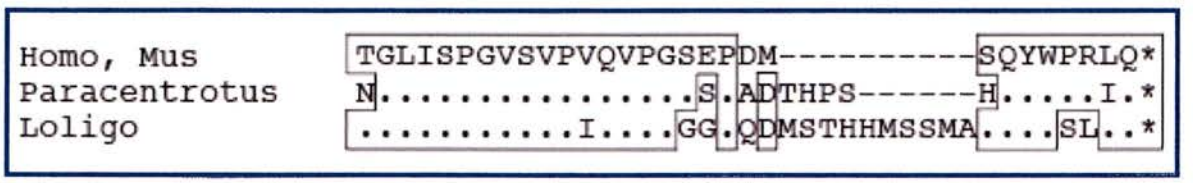



**Figure 1.19 A conserved C motif in the C terminus of PAX6.**

This shows a sequence comparison of the conserved motif found at the C terminus of the PAX6 protein in vertebrates, *Paracentrotus lividus* and *Loligo opalescens*. Dots represent identical amino acids, dashes represent introduced gaps and the asterisks represent the stop codon (Glardon et al., 1997).

This level of amino acid conservation suggests that the C terminal region of the PST domain is very important for the normal function of PAX6. The mutations described in section 1.7.6 all occur in this region and provide additional evidence for the functional importance of this part of the protein. However, at present this region of the PST domain is uncharacterised in terms of structure and function. Therefore, this set up the founding principles behind my PhD which were to characterise this C terminal region of PAX6 further and use it to investigate protein-protein interactions of PAX6.

## 1.9 Objectives.

### 1.9.1 Define the C terminal region using bioinformatics.

My main goal is to investigate the C terminal end of PAX6. However, as I have already mentioned this region is currently uncharacterised and so the first question that needs to be answered is "How should the C terminal region be defined?" This will be undertaken using bioinformatic programmes such as BLAST, which will be used to search for homology between the PST regions of PAX6 in human and other species, and JPRED, which will be used to investigate secondary structure in the C terminal region.

### 1.9.2 Create a series of clones for functional investigations.

On successful definition of the C terminal domain I will make clones that correspond to the C terminal domain, full length PAX6, and truncated forms of PAX6. I will also clone mutant copies of each construct that include the Q422R mutation, already mentioned. In addition I will make clones containing the 2 mutations that produce an elongated C terminal end of PAX6 (WHEDA and REMAR). Each peptide will be cloned into pDBLeu, a yeast 2-hybrid vector.

### 1.9.3 Identify interacting proteins using the yeast 2-hybrid system.

On successful completion of my cloning the next step will be to use the pDBLeu clones to investigate interactions with other proteins. This will be done by using the yeast 2-hybrid system to screen a mouse brain cDNA library.

### 1.9.4 Reconfirmation of the interactions.

The interactions identified from the yeast 2-hybrid screen would first be reconfirmed in the yeast 2-hybrid system by using pairwise interactions rather than library screening. Then I will try and finally reconfirm the interactions using an alternative method to the yeast 2-hybrid system, such as co-immunoprecipitation of tissue culture cells using tagged protein(s).

# Chapter 2 Materials and Methods

## 2.1 Bioinformatics.

### 2.1.1 BLAST (Basic Local Alignment Search Tool).

This is a program that searches for patches of similarity between inputted sequences and databases. BLASTP was used to identify regions of high conservation between PAX6 amino acid sequences and protein sequence databases (SPTR and TrEMBL). BLASTP searches were performed using the UK HGMP resource centre web page (http://menu.hgmp.mrc.ac.uk/cgi-bin/blast). Default settings were used for the searches except that the 'mask biased regions' filter was not used. This option filters out low compositional complexity regions. In the PAX6 PST region (used in the first BLAST search) it masks the region rich in proline, serine and threonine, preventing alignments in the PST domain. Hence by removing this filtering from the BLAST search it allows position-by-position alignment of the whole sequence.

BLASTN was also used (through http://www.ncbi.nlm.nih.gov/blast/) to find similar nucleotide sequences between query sequences, obtained from the sequencing of interactors pulled out of the yeast 2-hybrid screen, and nucleotide databases. Again default settings were used for this program.

BLAST2 Sequences (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html), another BLAST program was used to produces the alignment of two given sequences using the BLAST engine for local alignment. This program was mainly used in the comparison of human and mouse sequences of the 3 interesting interactors identified from the yeast 2-hybrid screens.

| Program | Use |
|---|---|
| BLASTP | Protein sequence used to search protein databases. |
| BLASTN | Nucleotide sequence used to search NCBI nucleotide databases. |
| BLAST 2 | Performs a comparison between 2 sequences (nucleotide or protein) using the BLAST algorithm. |

**Table 2.1 Differences between BLAST programs.**

### 2.1.2 CLUSTAL.

This is a program that performs multiple alignments of various PAX6 protein sequences. PAX6 sequences from different organisms, which were identified by BLAST searches on the PAX6 C terminus, were fetched from protein databases and compiled into a multiple sequence file. This multiple sequence file was then run through CLUSTAL to produce an alignment. The CLUSTAL alignment was run via the following web page http://www.hgmp.mrc.ac.uk /Registered/Option/clustal.html whereas protein sequences were retrieved from databases using a program called FETCH http://menu.hgmp.mrc.ac.uk/cgi-bin/fetchsequence.

### 2.1.3 JPRED.

This program predicts secondary structure for a protein sequence by combining a number of different prediction methods to form a consensus. The secondary structure prediction was carried out using the web page http://jura.ebi.ac.uk:8888 and the default settings were used for this program. (More detail of the specific use of this program will be explained in subsequent results sections).

### 2.1.4 Autoassembler™ 1.4.0.

This is a Mac based sequence analysis and manipulation program that was used to analyse sequencing results (see 2.7.3). This program was mainly used to compare sequencing results and also to produce alignments between similar stretches of sequence. This was particularly helpful when manipulating the sequence information of interesting interactors as it was easy to tell that certain isolated clones coded for the same protein.

### 2.1.5 EMBOSS (European Molecular Biology Open Software Suite).

A web interface to EMBOSS (http://www.hgmp.mrc.ac.uk/Registered/Webapp/emboss-w2h) was used to access sequence analysis software tools. Frequently used components were: Remap, which displays a sequence with protein translation and restriction sites indicated; Pepstats, which accurately calculates protein statistics such as molecular weight from protein sequence; and Matcher, which can compare 2 sequences and identify gaps or changes.

### 2.1.6 NCBI (National Centre for Biotechnology Information).

This website (http://www.ncbi.nlm.nih.gov/Entrez) was used to access many databases, which are outlined in the following table.

| Database | Notes |
|---|---|
| Nucleotide (Genbank) | Search for nucleotide sequences. |
| OMIM (Online Mendelian Inheritance in Man) | Investigate human genes and inherited diseases. |
| Protein | Search for protein sequences. |
| PubMed | Search for biomedical literature. |
| UniGene | Investigate gene-oriented clusters of transcript sequences. |

**Table 2.2 Summary of NCBI databases.**

## 2.2 DNA techniques.

### 2.2.1 Small scale preparation (miniprep) of plasmid DNA (QIAprep®).

**Buffer P1 (Resuspension Buffer)** 50mM Tris-Cl, pH 8.0;
10mM EDTA;
100μg/ml RNase A

**Buffer P2 (Lysis Buffer)** 0.2M NaOH,
1% w/v SDS

**Buffer N3 (Neutralization Buffer)** Recipe is proprietary and unavailable

**Buffer PE (Wash Buffer)** Recipe is proprietary and unavailable

**Buffer EB** 10mM Tris-Cl, pH 8.5

The following protocol is as outlined in the QIAprep Miniprep Handbook. All the centrifugation steps are 12 000x*g* at room temperature.

A single colony was used to inoculate 3ml of LB-broth, containing the appropriate selective antibiotic, and was incubated overnight at 37°C with vigorous shaking (200rpm). From this culture a 1.5ml aliquot was transferred into a 1.5ml microcentrifuge tube (Eppendorf™) and the cells were pelleted by centrifugation at 12 000x*g* for 1min. The supernatant was discarded and the pelleted bacterial cells were resuspended by vortexing in 250µl of Buffer P1. 250µl of Buffer P2 was then added to the resuspended cells and the tube gently inverted 6 times. This lysis reaction was not allowed to proceed for more than 5mins, then 350µl of ice-cold neutralising Buffer N3 was added and the tube immediately inverted 6 times. The mixture was centrifuged for 10mins, then the supernatant was pipetted into a preassembled QIAprep spin column in a 2ml collection tube. The plasmid DNA was bound to the column by centrifuging the supernatant through it for 1min. The column was then washed by adding 750µl of Buffer PE and centrifuged for 1min. The flow through was then discarded and the column completely dried by centrifuging for a further 1min. The plasmid DNA was then eluted into a clean eppendorf by adding 50µl of Buffer EB into the centre of the QIAprep column, allowing it to stand for 1min and then centrifuging for 1min. This eppendorf was then labelled and stored at -20°C.

### 2.2.2 Plasmid Purification (QIAGEN Plasmid Midi Protocol).

| | |
|---|---|
| **Buffer P1 (Resuspension Buffer)** | 50mM Tris-Cl, pH 8.0;<br>10mM EDTA;<br>100µg/ml RNase A |
| **Buffer P2 (Lysis Buffer)** | 0.2M NaOH, 1% w/v SDS |
| **Buffer P3 (Neutralization Buffer)** | 3M Potassium acetate pH 5.5 |
| **Buffer QBT (Equilibration Buffer)** | 750mM NaCl;<br>50mM MOPS, pH 7;<br>15% Isopropanol;<br>0.15% Triton® X-100 |
| **Buffer QC (Wash Buffer)** | 1M NaCl;<br>50mM MOPS, pH 7<br>15% Isopropanol |

| | |
|---|---|
| **Buffer QF (Elution Buffer)** | 1.25M NaCl;<br>50mM Tris-Cl, pH 8.5;<br>15% Isopropanol |
| **Buffer EB** | 10mM Tris-Cl, pH 8.5 |

The following protocol is as outlined in the QIAGEN® Plasmid Purification Handbook.

A single colony was picked from a freshly streaked selective plate and used to inoculate a 'day culture' of 5ml LB-broth, containing the appropriate selective antibiotic. This 'day culture' was incubated at 37°C for ~8hrs with vigorous shaking (200rpm). The 'day culture' was then diluted 1 in 1000 into 30ml of selective LB-broth and grown at 37°C overnight with vigorous shaking (200rpm). The bacterial cells were then harvested by centrifugation at 6000x*g* for 15mins at 4°C and the resulting pellet was resuspended in 4ml of Buffer P1 by vortexing. The resuspended cell pellet was then lysed, for 5mins at room temperature, by adding 4ml of Buffer P2. Then 4ml of ice-cold Buffer P3 was added, mixed by inverting 6 times and incubated for 15mins on ice, to allow precipitation of unwanted genomic DNA, protein, cell debris and SDS. This mixture was then centrifuged at 20 000x*g* for 30mins at 4°C. Whilst this centrifugation step took place, a QIAGEN-tip 100 was equilibrated by applying 4ml of Buffer QBT and allowing the column to empty by gravity flow. The resulting supernatant from the centrifuged mixture was filtered through a '597.5 Folded Filter' (Schleicher & Scuell), and then applied to the equilibrated QIAGEN-tip and allowed to enter the resin by gravity flow. The QIAGEN-tip was then washed twice with 10ml of Buffer QC and then the DNA was eluted with 5ml of Buffer QF into a glass tube (Corex). The DNA was precipitated by adding 3.5ml of room temperature isopropanol to the eluted DNA and then the precipitated DNA pelleted by centrifuging at 15 000x*g* for 30mins at 4°C. The supernatant was then carefully decanted and the pellet washed with 2ml of room temperature 70% ethanol. Again the DNA was pelleted by centrifugation at 15 000x*g* for 10mins at 4°C and the supernatant discarded. The pellet was allowed to air-dry for 5mins and then the DNA was redissolved in 200µl of Buffer EB. The resulting DNA was transferred to a clean eppendorf, which was then labelled and stored at -20°C

### 2.2.3 Spectrophotometric quantitation of DNA.

The concentration of DNA was determined by spectrophotometric measurement at 260nm ($OD_{260}$) using a UV Spectrophotometer (Ultrospec 3000, Pharmacia Biotech). The spectrophotometer was set to zero with $dH_2O$ and the optical density (OD) of a 1:200 dilution of DNA determined. This reading was then used to calculate the DNA concentration. An OD of 1 is equivalent to approximately 50μg/ml of double stranded DNA, 40μg/ml for single stranded DNA and 20μg/ml for single-stranded oligonucleotides. The DNA purity was determined by taking a further reading at 280nm; anything with an $OD_{260}/OD_{280}$ ratio of around 1.8 was considered acceptable. The formula required for calculating the concentration of genomic DNA is shown below.

$$\text{DNA concentration } \mu g/\mu l = \frac{(OD_{260})(50\mu g/ml)(\text{Dilution factor [D]})}{1000}$$

### 2.2.4 Restriction enzyme digestion of DNA.

Restriction endonuclease digestions were performed under the appropriate conditions as recommended by the manufacturer (Roche), and with the buffers supplied with the enzymes.

For diagnostic digests of miniprep DNA (see 2.2.1) 5μl of DNA (~2.5μg) was digested with 10 units of enzyme, in a total reaction volume of 20μl (made up with sterile water) and incubated at the optimal enzyme temperature (typically 37°C) for 1 hour.

For cloning strategies that required preparation of a vector fragment cut with a single enzyme, 5μg of genomic DNA was digested with 50 units of enzyme, in a total reaction volume of 60μl and incubated at the optimal enzyme temperature for 2 hours. If a double digest was required then 2 approaches were taken. If both enzymes worked effectively in the same reaction buffer then 50 units of each enzyme were combined with the DNA in a total reaction volume of 120μl and incubated at the optimal enzyme temperature for 2 hours. However, if the enzymes work in incompatible buffers then the DNA is digested, as for a single digest, and then the buffer conditions altered by adding suitable amounts of 500mM Tris and 500mM NaCl to mimic the preferred buffer conditions of the second enzyme. Once the buffer condition had been adjusted then 50 units of the second enzyme were added, the total reaction volume adjusted to 120μl and the digest incubated at the optimal temperature for a further 2 hours.

For cloning strategies that required preparation of an insert fragment, the restriction digest conditions were exactly the same as for preparing a vector fragment, however, 10μg of DNA were digested instead of 5μg.

### 2.2.5 5' dephosphorylation – *HK*™ Thermolabile Phosphatase.

*HK* Phosphatase (Epicentre Technologies) catalyses the dephosphorylation of protruding 5'-ends of DNA. This was used when trying to clone fragments into vector that had only been digested with a single restriction enzyme, to reduce self-ligation of the cut vector and allow insertion of the fragment. $CaCl_2$ was added to the 60μl vector digest to a final concentration of 5mM and incubated at 30°C for 5mins before the *HK* phosphatase was added. 5 units of *HK* phosphatase was then added and the reaction incubated at 30°C for 1 hour to dephosphorylate the DNA, and then at 65°C for 15mins to inactivate the *HK* phosphatase.

### 2.2.6 QIAquick™ PCR Purification Kit.

The QIAquick PCR purification kit (Qiagen) was used to remove restriction enzyme and restriction enzyme buffer from digested vector (see 2.2.4), as these may have interfered with subsequent ligation reactions. To each vector restriction digest reaction, 5 volumes of Buffer PB* was added to 1 volume of the reaction and mixed thoroughly before loading directly onto a QIAquick spin column. The cut DNA was bound to the column by centrifugation at 12 000xg for 1min, whilst the other reaction components passed through and were discarded. The DNA bound to the column was washed by adding 750μl of buffer PE followed by brief centrifugation and discarding the flow through. The columns were then dried by brief centrifugation. The cut DNA was eluted in 50μl of buffer EB, which was collected from the column into a fresh 1.5ml eppendorf during the final centrifugation.

*For components of QIAquick kit refer to the manufacturer.

### 2.2.7 Agarose gel electrophoresis.

| | |
|---|---|
| **Orange-G loading buffer** | 50% Glycerol, |
| **(5x Stock)** | 60mM EDTA, |
| | 0.25% Orange-G, |
| | 5x TBE |

| | |
|---|---|
| **Ethidium bromide (Bio-Rad)** | 10mg/ml in $dH_2O$<br>Store in darkness at room temperature (RT). |
| **Tris-borate (TBE)**<br>**(20x Stock, pH 8)** | 1.8M Tris,<br>1.8M Boric acid,<br>40mM EDTA |

DNA was visualised and size fractionated by agarose gel electrophoresis. The gels were prepared to concentrations ranging from 0.8 to 2% agarose (w/v) in the appropriate volume of 0.5x TBE buffer. The higher percentage gels were used to visualise low molecular weight DNA, and to compare small size differences. The low percentage gels were used to visualise high molecular weight DNA and to differentiate large size differences. The concentration of agarose was increased up to 4% to optimise resolution of smaller fragments such as PCR products, and in these cases NuSieve® GTG agarose (BioWhittaker Molecular Applications) was employed. Ethidium bromide was added at a concentration of 0.5μg/ml to molten agarose before each gel was poured. Once the gel had set it was submerged under buffer and DNA samples, combined with one-fifth volume Orange-G loading buffer, were loaded into the wells alongside an appropriate size marker. DNA size markers were *Hind*III cut bacteriophage λ (Gibco BRL, now Invitrogen) or *Hae*III cut ϕX174 (Promega) or DNA Molecular Weight Marker XVII (500bp ladder, Roche), depending on the size range of interest. The samples were then electrophoresed through the gel at 100 volts, until the DNA had run a sufficient distance. The DNA was then visualised by viewing under ultra violet light (on a transilluminator) and the image captured using the UVIdoc gel documentation system, produced by UVItec.

### 2.2.8 Recovery of DNA from agarose gels (QIAEX II).

The following protocol is as outlined in the QIAEX II Handbook and all solutions are as provided in the kit. All the centrifugation steps are 10 000x*g* at room temperature.

DNA was recovered from agarose using the QIAEX II Agarose Gel Extraction Kit (QIAGEN). DNA was first electrophoretically separated (see 2.2.7) and the target band was excised using a clean scalpel blade while briefly visualised by short wave UV transillumination. The excised band was weighed and 3 volumes of Buffer QX1 was added to 1 volume of gel for DNA fragments of 100bp to 4kb (for DNA fragments larger than 4kb an additional 2 volumes of $dH_20$ was also added). The QIAEX II beads were vortexed for 30secs, and then either 10μl or 30μl

QIAEX II beads was added to the sample. If the sample contained less than 2μg of DNA then 10μl of beads were added; alternatively if the sample contained more than 2μg of DNA then 30μl of beads were added. The sample was incubated at 50°C for 10mins, mixing every 2mins, to solubilise the agarose as well as allow the QIAEX II beads to bind the DNA. The sample was then centrifuged for 30secs and the supernatant was carefully removed and discarded. The pelleted QIAEX II beads were then washed once in 500μl of Buffer QX1 and then twice in 500μl of wash Buffer PE. After each wash the sample was centrifuged for 30secs and the supernatant was discarded. After the final wash step the pellet was left to air dry for 15mins until the pellet became white and then it was resuspended in 20μl of elution Buffer EB (10mM Tris-HCl, pH 8.5). The sample was incubated at room temperature for 5mins and then centrifuged for 30secs to remove the QIAEX II beads. The eluted DNA was pipetted into a clean eppendorf and stored at -20°C

### 2.2.9 Estimation of DNA concentration.

1-5μl of gel isolated DNA was made up to 10μl with 5x Orange G loading buffer and $dH_2O$ and loaded onto an agarose gel (see 2.2.7). After a short period of electrophoresis the relative amounts of vector and fragment DNA were estimated by comparison to 0.5μg of a relevant DNA size marker (see 2.2.7) under ultraviolet light visualisation.

### 2.2.10 Ligation of DNA.

| | |
|---|---|
| **T4 DNA ligase buffer (Roche)** | 660mM Tris-HCl, |
| **(10x Stock, pH7.5)** | 50mM $MgCl_2$, |
| | 50mM Dithiothreitol (DTT), |
| | 50mM ATP, |

Once digested with the appropriate restriction enzymes, and modified when necessary, DNA fragments were ligated together in 10μl reactions. A 3-fold molar excess of insert DNA: vector DNA was used for cloning experiments. Reactions were performed in 1x T4 DNA ligase buffer containing 1 unit of T4 DNA ligase at 4°C overnight. A vector only control and a vector only without T4 DNA ligase control were always set up to combat against incomplete restriction enzyme cutting. The completed ligation reactions were then stored at -20°C prior to transformation (see 2.6.1).

## 2.3 RNA techniques.

### 2.3.1 Tri Reagent™ preparation of RNA.

RNA was isolated from the cell lines H36CE2 and mv+ (see Table 2.8) using the Tri Reagent (Sigma) following the manufacturers protocol. All centrifugation was at 12 000x*g* at 4°C. In brief, a T75 flask of each cell type was grown until fully confluent (see 2.9.1) and 1ml of Tri Reagent was added directly to the flask to lyse the cells. The cell debris was then removed by spinning for 10mins and the supernatant was then combined with 200µl of chloroform. This mixture was then vortexed for 15secs, allowed to settle at room temperature for 5mins before spinning for a further 15mins. The upper aqueous phase contained the RNA and was transferred to a fresh eppendorf where it was combined with 500µl of isopropanol and stood at room temperature for 10mins. The mixture was then spun for 10mins and the RNA pellet washed with 1ml of 75% ethanol. The pellet was air dried for 5mins and then resuspended in 50µl of nuclease free water (Promega) and then stored in aliquots at -70°C. The RNA concentration was determined as for DNA (see 2.2.3) but by using this different formula:

$$\text{RNA concentration } \mu g/\mu l \;=\; \frac{(\text{OD}_{260})(33\mu g/ml)(\text{Dilution factor}\,[\text{D}])}{1000}$$

## 2.4 The polymerase chain reaction (PCR).

### 2.4.1 Oligonucleotide primer design.

Oligonucleotide primers were designed by eye and were between 17-21 bases in length. The software available on the Sigma-Genosys website (http://www.sigma-genosys.co.uk/index2.html) was used to check primer melting point and self-complementarity. Oligonucleotide primers were purchased from Sigma-Genosys UK Ltd. and were supplied as lyophilised products at a synthesis concentration of 0.3µM. The details of why different oligonucleotide primers were designed and their specific use in construct cloning will be discussed in more detail in the relevant results sections.

### 2.4.2 Polymerase chain reaction.

PCRs were carried out in a MJ Research Peltier Thermal Cycler (PTC-225 Gradient Cycler) in 0.5ml tubes, in a total reaction volume of 50µl made up with sterile $dH_2O$. The reaction mix contained 1x Mg-free Ampli*Taq* polymerase buffer (Perkin Elmer), 1.5mM $MgCl_2$, 200µM dNTPs (Pharmacia), 0.25µM each primer, 2.5 units of Ampli*Taq* polymerase (Perkin Elmer) and 10ng template DNA. After an initial denaturation step of 95°C for 30secs, the PCR conditions were 32 cycles of 94°C for 30secs, the appropriate primer annealing temperature for 30secs, and primer extension at 72°C for 30secs, and then a final extension stage at 72°C for 2mins. Details of the specific PCR primers used to generate various fragments can be seen in Table 2.3, which also gives details of the specific annealing temperature for primer pairs and very brief details of the use for each primer pair. The PCR products were always checked by running one tenth of the PCR reaction volume on an agarose gel. If the correct PCR product had been generated then it was stored at -20°C until used (see 2.4.8).

### 2.4.3 First strand cDNA synthesis.

First strand cDNA was amplified from the relevant RNA. The RNA for WHEDA and REMAR *PAX6* mutations had previously been prepared by Kathy Williamson from immortalised leukocyte cell lines from patients and I was instructed that 2µl of RNA would be sufficient to prepare first strand cDNA. RNA from lens epithelium cell lines had been prepared using the Tri Reagent method (Sigma), the extracted RNA had been quantified (see 2.3.1), and 2µg was used to prepare first strand cDNA. To the RNA 100pmol of random hexamer primers (pdN6), 10mM DTT and 1µl RNase inhibitor were added and made up to 14µl with nuclease free $H_2O$. The sample was then incubated at 70°C for 5mins. For each RNA sample 2 reactions were set up. In one (+AMV), the RNA was transcribed, whereas in the other, control sample, no reverse transcriptase was added (-AMV). To each reaction 1x Reaction Buffer, 1mM dNTPs and either 20 units AMV (Avian Myeloblastosis Virus) Reverse Transcriptase or $H_2O$ were added and incubated at 42°C for one hour and then 75°C for 8mins. The cDNA obtained was either placed on ice and used immediately for PCR, or stored at -20°C until use.

All reagents, apart from DTT, were from the First Strand cDNA Synthesis kit (Roche).

| Primer Pair | Sequence (5' → 3') | Annealing Temp | Frag Size | Information |
|---|---|---|---|---|
| ST001 | AAA AGT CGA CTG CCA GCA ACA CAC CTA GTC | 58°C | 464bp | Forward primer from the ASNT peptide. |
| ST005 | TTT TGC GGC CGC TTT TTA CTG TAA TCT TGG CCA GTA TTG | | | Reverse primer from *PAX6* stop codon. |
| ST002/ ST005 | AAA AGT CGA CTT CAT ACT CCT GCA TGC TGC | 58°C | 239bp | Forward primer from the SYSC peptide. |
| ST003/ ST005 | AAA AGT CGA CTA TGC AGA CAC ACA TGA ACA GTC | 58°C | 167bp | Forward primer from the MQTH peptide. |
| ST004/ ST005 | AAA AGT CGA CTA CCA CTT CAA CAG GAC TCA TT | 58°C | 125bp | Forward primer from the TTST peptide. |
| ST006 | TTT TGC GGC CGC TTT TTA CCG TAA TCT TGG CCA GTA TTG AG | 58°C | - | This reverse primer used instead of ST005 to produce Q422R mutant peptides. |
| ST007 | AAA AGT CGA CCA TGC AGA ACA GTC ACA GC | 58°C | 113bp | Forward and reverse primers used to produce missing portion of full length *PAX6*. |
| ST009 | TGA GCT AGC TCT ACA ATC TTC TG | | | |
| 9fwd | CAA AAT AGA TCT ACC TGA AGC | 53°C | 922bp | Designed by Kathy Williamson these primers amplified a specific region of WHEDA and REMAR mutant cDNA. |
| 13rvs | TAT CGA AGA CAC ACT CTA CC | | | |
| ST004/ ST013 | TTT TGC GGC CGC ATC CAT CCA GTC TAC ATT GTT C | 58°C | 437bp | Forward and reverse primers amplify from TTST peptide into the 3'UTR to produce WHEDA and REMAR peptides |
| ST015 | CCC ACA TAT GCA GAC ACA C | 58°C | 475bp | Forward and reverse primers to amplify a 3' portion of WHEDA to produce ASNT-WHEDA peptide by cloning. |
| ST031 | TTG CGG CCG CAT CCA TCC AGT CTA CAT TGT TC | | | |
| ST016 | AAG TCG ACC ATG TCC ACA GCC AGG GAA C | 52°C | 392bp | Forward and reverse primers to amplify N terminal region of *Homer3*. |
| ST017 | CTA GTG AAT TCT CCA CCA TC | | | |
| ST045 | AAG CGG CCG CTA TGC GAC TTC AGC TGA AG | 55°C | 417bp | Forward and reverse primers to amplify *PAX6*, which does not contain the paired domain. |
| ST046 | TAT TTG CCA TGG TGA AGC | | | |

**Table 2.3 PCR primers used in construct cloning.**

### 2.4.4 RT-PCR.

Each PCR contained 1x PCR Buffer*, 0.25μM forward and reverse primers, 200μM dNTPs*, 2.5 units of HotStar*Taq*™ DNA Polymerase* and 5μl of either the +AMV or –AMV cDNAs (see 2.4.3), in a total volume of 100μl. Primer sequences, annealing temperatures, and details of the use of each primer can be seen in Table 2.4. The PCR conditions were an initial denaturation step of 95°C for 15secs, followed by 35 cycles of 94°C for 30secs, the appropriate primer annealing temperature for 30secs, and primer extension at 72°C for 1min, and then a final extension stage at 72°C for 10mins. The resulting PCR products were then visualised by running one tenth of the PCR volume on an agarose gel.

* These reagents were from the HotStar*Taq* DNA Polymerase Kit (Qiagen).

### 2.4.5 Nested RT-PCR to generate WHEDA and REMAR *PAX6* mutants.

First strand cDNA was amplified from the relevant patient RNA (see 2.4.3). Nested PCR was then performed using a mix of Ampli*Taq* (Perkin Elmer) and *Pfu* (Stratagene) polymerase (9:1 unit ratio) to increase fidelity of the resulting fragment. The first set of PCR primers had been designed by Kathy Williamson (MRC HGU) to encompass exons 12 and 13 of *PAX6* (see 9fwd and 13rvs in Table 2.3). The second set of PCR primers were either ST004 and ST013 or ST015 and ST031 (see Table 2.3) and this generated specific fragments of mutated *PAX6*. The nested PCRs were setup as for normal PCR although the enzyme used was different. 1 unit of the Ampli*Taq*/*Pfu* mix was used per PCR and the PCR conditions were an initial denaturation step of 95°C for 30secs, and then 30 cycles of 94°C for 1min, the appropriate primer annealing temperature for 1min, and primer extension at 72°C for 1min, and then a final extension stage at 72°C for 10mins. As a control the nested PCRs were also set up using 2.5 units of Ampli*Taq* to check that the primers worked. Again the PCR products were always checked by running them out on an agarose gel and, then if correct, stored at -20°C until used (see 2.4.8).

| Primer | Sequence (5' → 3') | Melt Temp. | Frag Size. | Use |
|---|---|---|---|---|
| ST038 | CAG CCA AAA TAG ATC TAC CTG | 55°C | 600bp | Mouse forward *Pax6* RT-PCR primer. |
| ST039 | CGA TCA CAT GCT CTC TCC TT | | | Mouse reverse *Pax6* RT-PCR primer. |
| ST024 | CCC AGG TGG CTG TAG AGC | 55°C | 485bp | Mouse forward *Homer3* RT-PCR primer. |
| ST025 | CTC TAC ACA GTG CAA AGC TCA G | | | Mouse reverse *Homer3* RT-PCR primer. |
| ST026 | GTG CAG GAT GTG AAG CTG | 55°C | 609bp | Mouse forward *Trim11* RT-PCR primer. |
| ST027 | GCC TGC AGA TAG TCA TAG GG | | | Mouse reverse *Trim11* RT-PCR primer. |
| ST028 | CAA AAA TGC AGA CAT GTC G | 55°C | 485bp | Mouse forward *Dncl1* RT-PCR primer. |
| ST029 | CTA AGG GAG AAA AAA ATG GGG | | | Mouse reverse *Dncl1* RT-PCR primer. |
| Following primers were designed by Isabel Hanson. | | | | |
| IH076 | TGA GGA ATC AGA GAA GAC AGG | 58°C | 538bp | Human forward *PAX6* RT-PCR primer. |
| IH077 | GAA AGC TCA ACT GTT GTG TCC | | | Human reverse *PAX6* RT-PCR primer. |
| IH074 | CTC AGG CAG CTT CAG AGG | 58°C | 441bp | Human forward *HOMER3* RT-PCR primer. |
| IH075 | CCT TTC CAG GAC GAA TGG | | | Human reverse *HOMER3* RT-PCR primer. |
| IH078 | CCA GAA GTT GTG CCT ATG G | 58°C | 585bp | Human forward *TRIM11* RT-PCR primer. |
| IH079 | GCC GGC AGA TAG TCA TCG | | | Human reverse *TRIM11* RT-PCR primer. |
| IH081 | AAA AAT GCG GAC ATG TCG | 58°C | 492bp | Human forward *DNCL1* RT-PCR primer. |
| IH080 | TTT TGA GAG AAA ACA ACG TGG | | | Human reverse *DNCL1* RT-PCR primer. |
| GAPDH primers obtained from J. O'Donoghue and ATP5A1 primers obtained from B. Pickard. | | | | |
| GAPF | CAT CAC CAT CTT CCA GGA GC | 55°C | 450bp | Mouse forward *Gapdh* RT-PCR primer. |
| GAPR | ATG ACC TTG CCC ACA GCC TT | | | Mouse reverse *Gapdh* RT-PCR primer. |
| ATP5F | CAC AGC TGA GAT GTC CTC CA | 55°C | 415bp | Mouse forward *Atp5a1* RT-PCR primer. |
| ATP5R | CAC AGA GAT TCG GGG ATA A | | | Mouse reverse *Atp5a1* RT-PCR primer. |

**Table 2.4 RT-PCR Primers.**

### 2.4.6 Touchdown PCR.

The PCR is set up as for the RT-PCR, except that 200ng of template DNA was added to the reaction instead of +/- AMV cDNA (see 2.4.4). The primer sequences, melting temperatures (mp) and the details of the use of each primer can be seen in Table 2.5. The touchdown PCR was used to amplify MatchMaker Co-IP constructs as this type of PCR uses a cycling program with varying annealing temperatures that increases the specificity of the PCR product. The PCR conditions were an initial denaturation step of 95°C for 15mins, followed by 94°C for 30secs, 70°C for 30secs and 72°C for 1 or 2mins* which was repeated for 10 cycles with the annealing temperature reducing by 1°C each cycle so that a final 20 cycles of 94°C for 30secs, 60°C for 30secs and 72°C for 1/2mins could be performed. A final extension stage at 72°C for 10mins was also included. The resulting PCR products were then visualised by running one tenth of the PCR volume on an agarose gel (see 2.2.7).

* The extension time at 72°C was variable between 1 and 2mins due to the size of insert to be PCR'd. An extension time of 1min was used to generate PCR products of up to 1kilobase in length. To generate larger PCR products the extension time was increased by approximately 1min per kb DNA.

### 2.4.7 Poly(A) Reaction.

*Pfu* polymerase (Stratagene) exhibits a low error rate, making it the ideal choice for high-fidelity PCR amplification. However, PCR products that were generated using this proof reading polymerase were blunt ended, and so terminal deoxyadenosine needed to be added, in order that these PCR products could be inserted into the pGEM®-T cloning system. 5ng of PCR product, 1x Ampli*Taq* buffer, 0.2mM dATP and 5 units of Ampli*Taq* polymerase (Perkin Elmer) were incubated, in a total volume of 10µl, for 30mins at 70°C. The poly(A) PCR product was then ligated into pGEM-T using the normal protocol (see 2.4.8).



| Primer | Seq 5' → 3' | Tm (°C) | Use |
|---|---|---|---|
| ST048 | AAA ATT GTA ATA CGA CTC ACT ATA GGG CGA GCC GCC ACC **ATG** GAG GAG CAG AAG CTG ATC TCA GAG GAG GAC CTG GGT CAA AGA CAG TTG ACT GTA TCG | 63.6 | Forward primer adds a T7 promoter and a Myc antibody tag to 5' of pDBLeu constructs |
| ST047 | G TAA ATT TCT GGC AAG GTA GAC | 58.3 | pDBLeu reverse primer |
| ST049 | AAA ATT GTA ATA CGA CTC ACT ATA GGG CGA GCC GCC ACC **ATG** TAC CCA TAC GAC GTT CCA GAT TAC GCT CCA CCA AAC CCA AAA AAA GAG | 64.2 | Forward primer adds a T7 promoter and a HA antibody tag to 5' of pPC86 constructs |
| ST050 | G TAA ATT TCT GGC AAG GTA GAC | 58.3 | pPC86 reverse primer |

**Table 2.5 PCR primers for adding the T7 promoters and epitope tags.**

This table outlines the PCR primers that were designed to add a T7 promoter (purple box) and a cMyc epitope tag (blue box) for pDBLeu constructs or an HA epitope tag (orange box) for pPC86 constructs.

### 2.4.8 pGEM-T cloning strategy for PCR products (Promega).

**2x Rapid Ligation Buffer** 60mM Tris-HCl (pH 7.8),
20mM $MgCl_2$,
20mM DTT,
1mM ATP,
10% PEG.

The pGEM-T vector system was used as a convenient way of cloning PCR products. The single 3'-T overhang at the insertion site, which is added to this vector, is pivotal in producing a highly efficient PCR cloning system. Certain thermostable polymerases that were used, such as Ampli*Taq* (Perkin Elmer) and HotStar*Taq* (Qiagen), add a single deoxyadenosine to the 3'-end of the amplified fragments, which is able to ligate to the 3'-T overhang of the vector system. However, some PCR reactions that were performed required the use of a proofreading polymerase such as *Pfu* (Stratagene, see 2.4.5), which does not add 3'A's to the generated PCR product. Therefore, this has to be done artificially utilising a Poly(A) reaction (see 2.4.7), so that these PCR fragment can still be cloned using the pGEM-T system.

Once the PCR products have been generated so that they contain 3'A's then they can be cloned into the pGEM-T vector. The appropriate amount of PCR product (insert) to be included in the ligation reaction, was calculated using the following reaction:

$$\frac{ng\ of\ vector \times kb\ size\ of\ insert}{kb\ size\ of\ vector} \times (3:1)\ insert:vector\ molar\ ratio = ng\ of\ insert$$

The concentration of PCR product was estimated by running a tenth of the total volume of PCR product on an agarose gel (see 2.2.9). Then the correct amount of PCR product was added to 25ng of pGEM-T vector, 1x rapid ligation buffer and 3 units of T4 DNA ligase, made up to a total volume of 10µl with $dH_20$. The ligation was then incubated overnight at 4°C and then transformed into Library Efficiency™ DH5α™ Competent Cells (see 2.6.2) for blue/white screening.

## 2.5 Construct Cloning.

This section of the methods is by no means meant to explain the details of the cloning steps that were required to generate the constructs that I have used as part of my PhD. More detailed information of the specific cloning strategies can be seen in the construct cloning results chapter (Chapter 4).

### 2.5.1 GATEWAY Cloning.

| | |
|---|---|
| **Tris-EDTA (TE) Buffer, pH 8.0** | 10mM Tris-HCl, pH 8.0,<br>1mM EDTA,<br>Sterilise by autoclaving |

The GATEWAY cloning system (Invitrogen) was used to make *Homer3*, *Trim11* and *Dncl1* coupled to an N-terminal GST molecule. DNA of the appropriate pPC86 clones were cloned into the GATEWAY entry vector, pENTR3c. The LR Recombination Reaction was then used to transfer the appropriate pPC86 derived fragments from the pENTR3c clones into a GATEWAY destination vector pDEST™27.

1µl of pENTR3c clone miniprep DNA was combined with 300ng pDEST27, 1x LR Clonase™ Reaction Buffer and made up to 18µl with TE buffer pH 8. The LR Clonase Enzyme Mix was then thawed on ice and 2µl added to each reaction and mixed. After incubating at 25°C for 1hr, 1µl of Proteinase K was added to each reaction and incubated for 10mins at 37°C. 1µl of the LR recombination reaction was then transformed into 50µl Library Efficiency DH5α Cells (see 2.6.2).

### 2.5.2 Oligonucleotide annealing to make pDBLeu_ASNT-TTST.

This method was used to make a pDBLeu_ASNT_normal construct that was lacking the C terminal region of *PAX6*. A synthetic linker was created from 2 annealed oligos (see Table 2.6), which were designed to generate the complete *PAX6* clone without the C terminal domain. Details of this cloning procedure will be covered in more detail in the relevant results section.

| Primer Name | Sequence |
|---|---|
| ST018 (5'→3') | TAT GCA GAC ACA CAT GAA CAG TCA GCC AAT GGG CAC CTC GGG CTA GC |
| ST032 (3'→5') | A CGT CTG TGT GTA CTT GTC AGT CGG TTA CCC GTC GAG CCC GAT CGC CGG |
| | *Nde*I site *Not*I site |

**Table 2.6 Linker addition oligonucleotide pair used in cloning of pDBLeu_ASNT-TTST.**

Bases corresponding to *Nde*I and *Not*I restriction sites are coloured in red. An artificial stop codon (TAG) was also designed into the linker.

The oligos were ordered from Sigma-Genosys at 0.05μmol scale, with a 5' phosphate and PAGE purified, and were received as a lyophilised powder. Both complementary oligonucleotides were then resuspended at the same molar concentration of 100pmol/μl, using 1x T4 DNA ligase buffer. Then equal volumes (50μl) of both complementary oligos were mixed in a 1.5ml eppendorf tube, which was placed in a standard heatblock at 95°C. The heatblock was then removed from the apparatus and allowed to cool to 25°C (~45-60mins). The annealed oligos were then stored at 4°C until ready to use.

### 2.5.3 IMAGE clone for *Homer3*.

The *Homer3* clones that were isolated from the yeast 2-hybrid screen were found to be missing the N terminal 72 amino acids of the published sequence. Therefore, it was decided that it would be interesting to investigate whether full length *Homer3* protein behaved differently. Therefore, the *Homer3* image clone (IMAGE ID: 3602414) was ordered from the UK Human Genome Mapping Project Resource Centre website (http://www.hgmp.mrc.ac.uk/geneservice/reagents/products/image/index.shtml). The IMAGE clone arrived streaked on LB agar containing 50μg/ml ampicillin, and as suggested, the clone was restreaked as soon as possible onto the same medium and incubated overnight at 37°C to obtain single colonies. Then 10 isolated colonies of this restreak were grown up overnight at 37°C in the media suggested by HGMP (LB broth + $Mg^{2+}$ containing 50μg/ml ampicillin and 8% glycerol for subsequent freezing at -70°C). These colonies were then frozen on dry ice for 30mins and then transferred to a -70°C freezer for storage. A single loop of this glycerol stock was then used to seed a miniprep culture and the *Homer3* DNA was prepared using the Qiagen Miniprep Kit (see 2.2.1).

## 2.6 Transformation of bacterial cells with DNA.

**Antibiotic Selection - Ampicillin**

50mg/ml Ampicillin in $dH_2O$,
Filter sterilise,
Store at -20°C.
Ampicillin used at 50μg/ml,

**Antibiotic Selection - Kanamycin**

25mg/ml Kanamycin in $dH_2O$,
Filter sterilise,
Store at -20°C.
Kanamycin used at 25μg/ml

**Isopropyl-β-D-thiogalactopyranoside (IPTG)**

100mM in $dH_2O$
Filter sterilize to 0.22 microns,
Store at -20$^0$C

**Luria-Bertani (LB)-agar**

1.5% w/v Bacto-agar,
1% w/v Bacto-tryptone,
0.5% Bacto-yeast extract,
1% w/v NaCl,
pH to 7.2 with NaOH.
Sterilise by autoclaving.

**Luria-Bertani (LB)-broth**

1% w/v Bacto-tryptone,
0.5% Bacto-yeast extract,
0.5% w/v NaCl,
pH to 7.2 with NaOH.
Sterilise by autoclaving.

**Luria-Bertani (LB)-broth + $Mg^{2+}$**

1% w/v Bacto-tryptone,
0.5% Bacto-yeast extract,
0.5% w/v NaCl,
0.246% w/v $MgSO_4$,
0.1% w/v Glucose,
pH to 7.2 with NaOH, sterilise by autoclaving.

| | |
|---|---|
| **SOC Replacement Medium** | 0.4% Glucose in LB-broth + $Mg^{2+}$,<br>Mixture filter sterilised. |
| **5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)** | 20mg/ml in Dimethylformamide<br>Store in darkness at -20°C. |
| **X-Gal LB-Agar Plates.** | LB-Agar plate containing the appropriate selection media.<br>100µl IPTG,<br>40µl X-Gal,<br>Allow the IPTG and X-Gal to diffuse into the agar for at least 30mins. |

Library Efficiency DH5α Competent Cells (Invitrogen) and XL1-Blue Subcloning-Grade Competent Cells (Stratagene) were transformed with DNA by heat shock, according to the manufacturer's instructions. Aliquots of competent cells were stored at -70°C prior to use.

### 2.6.1 XL1-Blue Subcloning-Grade Competent Cells.

These cells were used for 'general-purpose' transformations and for subcloning steps. Sufficient eppendorf tubes for the required number of transformations were prechilled on ice. An aliquot of competent cells was thawed on ice and 50-100µl dispensed into each tube. Next either 10ng of supercoiled plasmid DNA or 5% of the volume of cells for ligation reactions (e.g. for 50µl cells add 2.5µl of ligation reaction), was added to each tube, swirled to mix and the tubes left on ice for 20mins. Also for every transformation that was performed a positive control was always included; 0.1ng of the pUC18 control plasmid was added to a 50-100µl aliquot of cells. After incubation on ice the cells were then heat-pulsed in a 42°C water bath for exactly 45secs followed by incubation on ice for 2mins. 0.9ml of SOC replacement medium were added and the tubes incubated with shaking at 37°C for 1 hour. 10µl and 200µl aliquots of the transformation mix were plated onto fresh LB-agar plates containing the appropriate selection media and cultured at 37°C overnight.

### 2.6.2 Library Efficiency DH5α Competent Cells.

These cells were used when blue/white screening was required (e.g. when using pGEM-T cloning of PCR generated fragments, see 2.4.8). Sufficient eppendorf tubes for the required number of transformations were prechilled on ice. An aliquot of competent cells was thawed on ice and 50μl dispensed into each tube. Next 2.5μl (5% of the volume of cells) of the pGEM-T ligation reaction was added to each tube, swirled to mix and the tubes left on ice for 30mins. Also for every transformation that was performed a positive control was always included; 500pg of the pUC19 control plasmid was added to a 50μl aliquot of cells. After incubation on ice the cells were then heat-pulsed in a 37°C water bath for exactly 20secs followed by incubation on ice for 2mins. 0.9ml of SOC replacement medium were added and the tubes incubated with shaking at 37°C for 1 hour. 5μl and 200μl aliquots of the transformation mix were plated onto fresh X-Gal LB-agar plates and cultured at 37°C overnight.

### 2.6.3 ElectroMAX™ DH10B™ Cells.

Isolated DNA of positive interactors was electroporated into ElectroMAX DH10B cells (GibcoBRL) using the following protocol. 1μl of DNA (c. 50ng) was added to the 40μl aliquot of competent cells and cooled on ice for 5mins. A 10pg pUC19 positive control was also included in every transformation to estimate transformation efficiency. The gene-pulser apparatus was then set to a capacitance of 25μF, voltage of 2.5kV and impedence of 100Ω. The cells were then placed in a chilled 0.2ml electroporation cuvette and pulsed with a time constant of 4.5msec. The cuvette was removed from the gene pulser and 1ml of SOC replacement medium was immediately mixed with the cells; they were then cultured at 37°C for one hour. The cells were then spun down at 10 000x*g*, the cell pellet resuspended in 150μl of SOC replacement medium and this was then plated onto fresh LB-agar plates containing 50μg/ml of ampicillin and cultured at 37°C for 16 hours. This will selectively isolate the pPC86-Y plasmid, which is ampicillin resistant, and not the pDBLeu-X plasmid.

### 2.6.4 Restreaking bacterial colonies.

Short-term storage of bacterial transformants is achieved by restreaking positive colonies on LB agar plates containing the correct antibiotic selection. Transformants can be stored by this method for up to 4 weeks at 4°C.

A single positive clone is streaked onto an LB agar plate containing the correct antibiotic selection and the plate is incubated overnight at 37°C. The resulting plate of single colonies can then be stored at 4°C.

### 2.6.5 Glycerol stocks of bacterial colonies.

**Glycerol Solution** 15% Glycerol in LB-broth + $Mg^{2+}$,

Long-term storage of bacterial transformants is attained by resuspending an aliquot of grown up cells in 15% glycerol in selective media. These transformants can be stored indefinitely at -70°C.

A 1ml aliquot of overnight bacterial culture is centrifuged at 10 000x*g* for 1min. The cell pellet is resuspended in 1ml glycerol solution and then flash frozen on dry ice for 30mins. The frozen glycerol stock is then transferred to -70°C.

## 2.7 Sequencing.

DNA was sequenced using the BigDye™ terminator sequencing kit (PE Applied Biosystems), which is a variation of the Sanger end terminator sequencing technique (Sanger *et al.*, 1977).

### 2.7.1 ABI Prism® BigDye Terminator Cycle Sequencing Kit.

**Termination Ready Reaction Mix**

Dye terminators:
ddATP labelled with dichloro[R6G],
ddCTP dichloro[ROX],
ddGTP dichloro[R110]
ddTTP dichloro[TAMRA]
Deoxynucleoside triphosphates
(dATP, dCTP, dITP, dUTP)
Ampli*Taq* DNA polymerase, FS,
with thermally stable pyrophosphatase
$MgCl_2$
Tris-HCl buffer, pH 9.0

The sequencing reactions were set up in 0.5ml PCR tubes and contained: 3.2pmol of sequencing primer, approximately 750ng of template DNA, 4μl of BigDye terminator ready reaction mix and deionised water to a total reaction volume of 20μl. Each reaction was mixed well and collected by brief centrifugation. The sequencing process was carried out on a MJ Research Peltier Thermal Cycler (PTC-225 Gradient Cycler), and comprised of 25 cycles of 96°C for 30secs, 47°C for 15secs (this melting temperature varies for different sequencing primers - see Table 2.7) and 60°C for 4mins. The sequencing products were then purified.

### 2.7.2 Purification of BigDye sequencing products.

The sequencing reactions were purified using Edge Gel Filtration Cartridges (Edge BioSystems). The Edge Gel Filtration Cartridges were prepared by spinning for 2mins at 750xg and then placing in a fresh 1.5ml eppendorf. The sequencing reaction was then applied to the centre of the packed column, the cap closed and spun for 2mins at 750x*g*. The eluate was then dried for 15mins in a DNA Speed Vac® (Savant) set at medium drying rate. These dried samples were then stored at -20°C until sequenced by the sequencing service.

### 2.7.3 Sequencing Service.

Sequencing gels were run either by Angie Fawkes or Alison Condie on an ABI 3100 Capillary Sequence Analyser. The resulting sequence traces were then analysed using either the Chromas program (Technelysium – PC based) or the Autoassembler program (Perkin-Elmer - Apple based).

| Primer Name | Sequence (5' → 3') | Melting Temp. | Use |
|---|---|---|---|
| T7 | ATT ATG CTG AGT GAT ATC CC | 47°C | Sequence fragments cloned into pGEM-T cloning system. |
| SP6 | ATT TAG GTG ACA CTA TAG AA | 47°C | Sequence fragments cloned into pGEM-T cloning system. |
| ST014 | GAT CAC TCG TTT CAG CTG | 56°C | Check 3' end of *Homer3* clones. |
| pPC86-F | TAT AAC GCG TTT GGA ATC ACT | 56°C | Sequence positive clones from yeast 2-hybrid screening. |
| pPC86-R | GTA AAT TTC TGG CAA GGT AGA C | 56°C | Sequence positive clones from yeast 2-hybrid screening. |
| ST033 | GAC CAA GGA CCA GGA GAT | 56°C | To check full sequence of *Homer3* cDNA. |
| ST034 | CTC AGT GTT TCC TGT CCC | 56°C | To check full sequence of *Homer3* cDNA. |
| ST035 | GGA GGC TGG AGA AGT CAC | 56°C | To check full sequence of *Triml1* cDNA. |
| ST036 | AGC TTT CGG CAC TAA GTG | 56°C | To check full sequence of *Triml1* cDNA. |
| ST037 | AGT GTC TGC CCC ACT TTC | 56°C | Sequence N terminal region of full length *Homer3*. |

**Table 2.7 Sequencing Primers used to check or identify various clones.**

## 2.8 Yeast 2-hybrid System.

The yeast 2-hybrid work was carried out using the ProQuest™ Two-Hybrid System (Gibco BRL). The system was used to investigate possible interactions between the proteins encoded by a mouse brain cDNA library (Gibco BRL) and the protein generated by the construct pDBLeu_TTST_norm or pDBLeu_ASNT_norm, which correspond to the wild type C terminal domain of PAX6, and the PST region of PAX6 respectively, in the yeast 2-hybrid vector pDBLeu.

### 2.8.1 ProQuest Two-Hybrid Mouse Brain cDNA Library.

The cDNA library was used to prepare double-stranded plasmid DNA that was used with the ProQuest Two-Hybrid System. The Gibco BRL protocol for the amplification and preparation of plasmid DNA from an amplified cDNA library was followed exactly and the resulting DNA was stored at -20°C in 1μg/μl aliquots.

### 2.8.2 Small-scale preparation of competent MaV203 cells.

| | |
|---|---|
| **Tris-EDTA (TE) Solution (10x)** | 100mM Tris-HCl pH 7.5,<br>10mM EDTA,<br>Sterilise by autoclaving |
| **Lithium Acetate (LiAc) Solution (10x)** | 1M Lithium acetate,<br>Sterilise by autoclaving. |
| **TE/LiAc Solution (1x)** | 10% v/v 10x TE,<br>10% v/v 10x LiAc,<br>Make fresh and store on ice. |
| **YPAD broth** | 1% w/v Bacto-yeast extract,<br>2% w/v Bacto-peptone,<br>2% w/v Glucose,<br>0.01% w/v Adenine sulphate,<br>pH to 6 with HCl,<br>Sterilise by autoclaving. |

| | |
|---|---|
| **YPAD Agar** | 2% w/v Bacto-agar,<br>non-autoclaved YPAD medium,<br>pH to 6 with HCl,<br>Sterilise by autoclaving. |

MaV203 cells were streaked out from the glycerol stock (supplied in Gibco BRL Two-Hybrid System Kit) onto a YPAD plate and grown for 48 hours at 30°C. Several of the resulting single colonies were then resuspended in 50µl of $dH_2O$ and spread in a circle into the middle of a fresh YPAD plate, which was then grown overnight at 30°C. The resulting circle of cells was then resuspended in 5ml $dH_2O$ and the optical density at 600nm was determined. A sufficient amount of this cell suspension was added to 100ml YPAD broth in a 500ml flask to produce a final $OD_{600}$ of 0.1. The MaV203 culture was then incubated at 30°C with shaking (250rpm) for 4-5 hours until the $OD_{600}$ reached 0.4. The 100ml culture was split into 2x 50ml aliquots and the yeast cells of each aliquot were then pelleted by spinning at 3000x*g* for 5mins at room temperature. The cell pellet was then resuspended in 20ml $dH_2O$ and spun for a further 5mins at 3000x*g* at room temperature. The resulting pellet was washed in 10ml of 1x TE/LiAc solution and spun for a further 5mins at 3000x*g*. The final cell pellet was then suspended in 175µl of 1x TE/LiAc solution and the two aliquots were pooled. This protocol produced enough MaV203 cells for 7 yeast transformations (see 2.8.3).

### 2.8.3 Small-scale yeast transformation.

| | |
|---|---|
| **Amino Acid Mix** | 2g of each of the following:<br>Adenine sulphate, alanine, arginine, aspartic acid, asparagine, cysteine, glutamic acid, glutamine, glycine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, valine. |
| **PEG/LiAc Solution** | 10% v/v 10x TE,<br>10% v/v 10x LiAc,<br>80% v/v 50% PEG-3350,<br>Make fresh and store on ice. |

| | |
|---|---|
| **SC Agar** | 0.705% w/v Difco™ Yeast Nitrogen Base w/o amino acids,<br>0.142% w/v Amino Acid Mix,<br>pH to 5.9 with NaOH,<br>2% w/v Bacto-agar,<br>Autoclave to sterilise and melt agar. |
| **SC-Leu Agar Plates** | 2% w/v Glucose,<br>0.16mM Uracil,<br>0.8mM Histidine,<br>0.32mM Tryptophan,<br>Added to autoclaved SC Agar, mixed and poured. |
| **SC-Leu-Trp Agar Plates** | 2% w/v Glucose,<br>0.16mM Uracil,<br>0.8mM Histidine,<br>Added to autoclaved SC Agar, mixed and poured. |
| **Sonicated Salmon Sperm** | 50µl Sonicated salmon sperm (10mg/ml),<br>Boiled for 5mins,<br>Incubate on ice until used. |

For each transformation 50µl of MaV203 cells were combined with 5µl of freshly boiled carrier DNA (sonicated salmon sperm) and 100ng of each plasmid DNA in an autoclaved 1.5ml eppendorf. This was mixed gently by pipetting up and down and then 300µl of PEG/LiAc solution was added, mixed and then incubated at 30°C for 30mins. The yeast cells were then 'heat-shocked' for 15mins at 42°C, centrifuged at 7000x*g* for 30secs and the cell pellet was resuspended in 500µl of $dH_2O$. To obtain well-separated colonies, 10µl and 200µl of the suspension was spread over the appropriate SC agar selection plates (SC-Leu plates for single transformants or SC-Leu-Trp for double transformants). The plates were incubated at 30°C for 48-72 hours until colonies appeared.

### 2.8.4 Determining the optimum 3AT concentration needed for library screening.

*HIS3* is a reporter gene that is used as the first selection for possible interacting proteins. The enzyme produced by this reporter gene (imidazole glycerol phosphate dehydratase) is inhibited by 3-amino-1, 2, 4-triazole (3AT) in a dose-dependent manner. In order for the library screening to be a success any background growth, due to self-activation of the *HIS3* reporter by pDBLeu-X, has to be eliminated so that only positive interactors will be seen. Therefore, following the protocol outlined by the ProQuest 2-Hybrid System (Gibco BRL, pages 16-17), pDBLeu_TTST_norm and pDBLeu_ASNT_norm were titrated against 3AT. It was found that 7.5mM 3AT was needed to eliminate basal levels of *HIS3* reporter gene expression for pDBLeu_TTST_norm and 10mM 3AT was needed for pDBLeu_ASNT_norm.

### 2.8.5 Screening the cDNA library for possible interactions with pDBLeu constructs.

| | |
|---|---|
| **SC-Leu-Trp-His + 3AT Agar Plates** | 2% w/v Glucose,<br>0.16mM Uracil,<br>7.5 or 10mM 3-Amino-1,2,4-Triazole (3AT),<br>Added to autoclaved SC Agar and incubated at 50°C for 30mins before pouring into plates. |

*PAX6* constructs and the mouse brain cDNA library were co-transformed into library efficiency MaV203 competent cells (Gibco BRL) using the method in the ProQuest 2-Hybrid protocol (Gibco BRL). Firstly, 2x 250µl aliquots of cells were each combined with 7.5µg of DNA binding domain vector (pDBLeu) containing a test gene (i.e. pDBLeu_TTST_norm or pDBLeu_ASNT_norm) and 7.5µg of activation domain vector (pPC86) containing the cDNA library (see 2.8.1) in a 1.5ml eppendorf. Each eppendorf of cells and DNA was combined with 1.5ml of the supplied PEG/LiAc solution and incubated at 30°C for 30mins. Then 88µl of DMSO were added to each tube and the cells were heat shocked for 20mins at 42°C. The transformed cells were then pelleted at 640x*g* for 5mins and each cell pellet was resuspended in 8ml of 0.9% NaCl. To ensure that sufficient cells were transformed to represent the cDNA library several dilutions (ranging from 1:50 to 1:1500) of the transformed cells were plated onto 10cm SC-Leu-Trp plates and incubated at 30°C for 3-4 days. The remainder of the transformed cells were then plated in 400µl aliquots onto 40x 15cm SC-Leu-Trp-His+3AT plates and

incubated at 30°C for 3 days (correct concentrations of 3AT determined in 2.8.4). In tandem with this library screen a control transformation was also included to ensure that the transformation worked correctly. This control involved combining 25µl cells with 1µg of pMAB12 and pMAB37 DNA and incubating with 180µl PEG/LiAC at 30°C for 30mins. Then 10.8µl of DMSO were added and the cells were heat shocked for 20mins at 42°C. The cells were then pelleted as before and resuspended in 1ml 0.9% NaCl. Then 1:100 and 1:1000 dilutions of the transformed cells were plated onto 10cm SC-Leu-Trp plates and incubated at 30°C for 3 days.

After the 3 days incubation at 30°C, both the screen plates and the control plates were replica cleaned, using velvet and then returned to 30°C for further incubation. The development of *HIS3* positive colonies was then recorded for up to a week at which time *HIS3* positive colonies were stored as described in 2.8.6.

### 2.8.6 Storing *HIS3* positive transformants.

**Yeast Freeze Solution**

40% v/v Glycerol,
2% w/v Glucose,
0.16mM Uracil,
0.8mM Histidine,
In SC media (SC Agar without Bacto-agar).

Before testing *HIS3* positive transformants further, it was vital that these colonies were stored so that re-testing could be accomplished. This was done using this derived technique. A single colony that grew on the SC-Leu-Trp-His+3AT plate was dissolved in 50µl of $dH_2O$ and then spread on a quadrant of an SC-Leu-Trp plate and grown at 30°C for 3 days. This whole quadrant was then scraped off the plate and dissolved in 1ml of yeast freeze solution and stored at -80°C.

### 2.8.7 Confirming *HIS3* positive transformants using 3 reporter genes.

| | |
|---|---|
| **SC-Leu-Trp-Ura Agar Plates** | 2% w/v Glucose,<br>0.8mM Histidine,<br>Added to autoclaved SC Agar, mixed and poured. |
| **SC-Leu-Trp + 5FOA Agar Plates** | 2% w/v Glucose,<br>0.16mM Uracil,<br>0.8mM Histidine,<br>0.2% w/v 5-Fluoroorotic acid (5FOA),<br>Added to autoclaved SC Agar and incubated at 50°C for 30mins before pouring into plates. |

The resulting colonies were further tested to confirm whether any interaction was positive using 3 reporter genes (*HIS3*, *URA3* and *lacZ*) that have been integrated into the yeast genome. Technical information for this can be found in the ProQuest 2-Hybrid system manual (Gibco BRL, pages 19-22).

*HIS3*: As already mentioned positive interaction means that growth will occur on plates lacking histidine. As before plates contained the optimal concentration of 3AT to eliminate background growth.

*URA3*: As with *HIS3*, a positive interaction will allow growth on plates lacking uracil. This tests for high-level activation of the *URA3* reporter gene and so will identify strong interactions. However, the *URA3* reporter gene can be used for additional confirmation of positive interactions as the induction of *URA3* causes 5-fluoroorotic acid (5FOA) to be converted into toxic 5-fluorouracil. Therefore, the growth of the cells containing interacting proteins will be inhibited on plates containing 5FOA. In contrast this tests for low-level activation of the *URA3* reporter gene and is therefore able to test for weak interactions.

*LacZ*: Interacting proteins will activate the *lacZ* reporter gene, which will cause X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) to turn blue in a β-gal assay.

For this method, *HIS3* positive transformants were prepared as single colonies from the stored glycerol stocks, by restreaking the glycerol onto an SC-Leu-Trp plate quadrant and incubating it at 30°C for 3 days. Single colonies were then picked and plated in quadruplicate onto a single plate called a 'master plate' (see Figure 2.1). This master plate also included the 5 supplied yeast control strains (A-E), which were plated in duplicate and also the relevant pDBLeu-X + pPC86 negative control, which was also plated in duplicate.

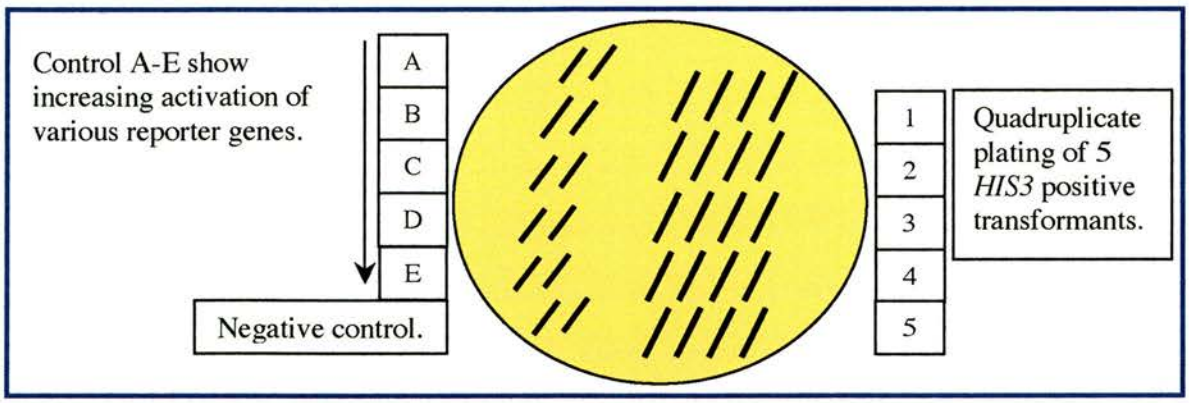



**Figure 2.1 Yeast 2-Hybrid master control plate layout.**

The master plate was incubated at 30°C for 24 hours. The master plate was then replica plated, using velvet, onto the following plates, in the order listed: YPAD containing a Hybond™-N nylon membrane (Amersham pharmacia biotech); SC-Leu-Trp-Ura; SC-Leu-Trp-His + 3AT (replica cleaned); SC-Leu-Trp + 5FOA (replica cleaned). These plates were incubated at 30°C for 24 hours and then the YPAD plate was used in an X-Gal assay (see 2.8.8) and the remaining plates were replica cleaned and then further incubated at 30°C for a further 48 hours. The resulting phenotypes of the potential positives were then compared to the Yeast Control Strains A-E and only the colonies that induced all three reporter genes were further analysed.

### 2.8.8 X-Gal Assay.

| | |
|---|---|
| **5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)** | 100mg/ml in Dimethylformamide<br>Store in darkness at -20$^0$C. |
| **Z Buffer** | 60mM Anhydrous di-sodium hydrogen orthophosphate ($Na_2HPO_4$),<br>40mM Anhydrous sodium phosphate ($NaH_2PO_4$),<br>10mM Potassium chloride (KCl),<br>1mM Anhydrous magnesium sulphate ($MgSO_4$),<br>Adjust to pH 7.0 and autoclave. |
| **Z/X-Gal Reaction Buffer** | 100µl X-Gal (100mg/ml),<br>60µl β-mercaptoethanol,<br>10ml Z buffer,<br>This is enough for one membrane. |

As described in the previous section (see 2.8.7) the master plate had been replica plated directly onto a Hybond-N nylon membrane and incubated for between 18-24 hours at 30°C. Two round 125mm Whatman 541 filter papers were stacked in a 15cm Petri dish and saturated with 8ml of the Z/X-Gal Reaction Buffer. Any air bubbles were removed as well as any excess reaction buffer. The Hybond-N membrane was flash frozen for 30secs in liquid nitrogen to permeabilise the colonies, and then left to thaw at room temperature. The membrane filter was then transferred to the prepared 15cm Petri dish, colony-side up and incubated at 37°C (Parafilm was wrapped around the Petri dish to minimise the smell of β-mercaptoethanol). The appearance of the blue colour was monitored over 24 hours and then finally scored after 24 hours according to the rate at which the colonies turned blue: strong interactors show blue colour within 1 hour (e.g. Yeast Control Strain E); weak interactors show blue colour within 24 hours (e.g. Yeast Control Strain B). The appearance of blue colonies in yeast cells transformed with only one of the yeast expression plasmids would indicate that the expressed fusion protein auto-activates the system.

### 2.8.9 Plasmid DNA extraction from Yeast.

**SC-Leu-Trp media** 2% w/v Glucose,
0.16mM Uracil,
0.8mM Histidine,
In SC media (SC Agar without Bacto-agar).

Plasmid DNA isolated from yeast is typically not suitable for restriction analysis or sequencing. Therefore, plasmid DNA was first isolated from yeast, and then introduced into *E. coli* by electroporation (see 2.6.3). Then the resulting transformants were miniprepped (see 2.2.1), and this DNA was suitable for restriction analysis and sequencing.

The method used for the isolation of plasmid DNA from yeast was as described in the ProQuest 2-Hybrid System manual (Gibco BRL, pages 41-42). For this method single colonies were needed and so the glycerol stock (see 2.8.6) was streaked out onto a quadrant of an SC-Leu-Trp plate and incubated at 30°C for 3 days. A single colony was then suspended in 1ml of SC-Leu-Trp media and incubated at 30°C with shaking for 24 hours. The resulting culture was spun at 14 000x*g* for 30secs and the yeast cell pellet was suspended in 100µl of freshly prepared 3% sodium dodecyl sulphate (SDS), 0.2M sodium hydroxide (NaOH). This was incubated at room temperature for 15mins and then 500µl of 1xTE, 60µl of 3M sodium acetate and 600µl of phenol: chloroform: isoamyl alcohol (25:24:1, Fluka) were added in series mixing after each

addition. The resulting mixture was then vortexed at full speed for 2mins and then centrifuged at 14 000x*g* for 2mins. The upper aqueous phase was transferred to a new 1.5ml eppendorf and another 600µl of phenol: chloroform: isoamyl alcohol was added, vortexed and spun as before. The upper aqueous phase was once again transferred to a fresh 1.5ml eppendorf and 650µl of cold isopropanol was added, mixed and incubated at -20°C for 20mins. The tube was then centrifuged for 5mins at 14 000x*g*, the supernatant carefully discarded, and the pellet washed in 100µl of 70% ethanol. After spinning for a further 5mins at 14 000x*g* and removing the supernatant the pellet was suspended in 10µl of 1xTE and stored at -20°C until used to electroporate ElectroMAX DH10B cells (see 2.6.3).

## 2.9 Cell Culture.

| | |
|---|---|
| **DMEM – Full media** | 90% (v/v) DMEM (Dulbecco's Modified Eagle Medium),<br>10% (v/v) FCS (Foetal Calf Serum),<br>2mM Glutamine,<br>100units/ml Penicillin,<br>100µg/ml Streptomysin,<br>Store at 4°C but warm to 37°C before use. |
| **DMEM – No antibiotics** | 90% (v/v) DMEM (Dulbecco's Modified Eagle Medium),<br>10% (v/v) FCS (Foetal Calf Serum),<br>2mM Glutamine,<br>Store at 4°C but warm to 37°C before use. |
| **Freeze Buffer** | 10% (v/v) DMSO,<br>90% (v/v) FCS, Store at -20°C. |
| **Trypsin: Versene (T:V)** | 1:1 Trypsin: Versene,<br>Store at 4°C but warm to 37°C before use. |
| **V3 Media** | 1:3 Versene: PBS<br>Store at 4°C but warm to 37°C before use. |

### 2.9.1 Growth of cell lines *in vitro.*

An Envair Class II Microbiological Safety Cabinet provided the sterile environment for tissue culture. Cells were grown in T75 tissue-culture flasks with filter caps (CellStar®) in DMEM full medium at 37°C with 5% carbon dioxide. The 2 cell lines that were commonly used are outlined in Table 2.8. Adherent cell layers were routinely passaged at subconfluency, normally every 2-3 days

| Cell Line | Info | Supplied by |
|---|---|---|
| mv+ | Mouse lens epithelium cell line | Veronica van Heyningen. |
| H36CE2 | Human lens epithelium cell line | Veronica van Heyningen. |

**Table 2.8 Cell line information.**

### 2.9.2 Passage of surface-adherent cells.

The growth medium was aspirated from the flask and the cells were washed once in 12ml of V3 media. Then 12ml of T:V was added to the cells and the flask was incubated at 37°C until the cells had begun to dissociate from the flask, typically 10-15mins. The detached cells were then re-seeded at a suitable density, typically 1:10 or 1:20, into a fresh T75 flask containing 30ml pre-warmed DMEM full media. The flask was then returned to the 37°C incubator.

### 2.9.3 Preparation and recovery of cell stocks.

For long-term storage cells were stored in liquid nitrogen. The cells were first treated as if they were to be normally passaged. However, once the cells had been detached from the flask, they were transferred to a falcon tube and the volume made up to 50ml with DMEM full medium. The cell suspension was then centrifuged for 3mins at 750x*g*. The supernatant was carefully removed and the cell pellet was resuspended in 9ml of Freeze Buffer. 1ml of cell suspension per vial was aliquoted into NUNC™ cryotubes (Nalge Nunc International) and the cells slow-frozen at -70°C before transferring to liquid nitrogen the next day. To recover the cells, a cryotube was removed from liquid nitrogen, defrosted, and in a safety cabinet, used to seed a T75 flask containing 30ml of pre-warmed DMEM full medium. Following overnight incubation at 37°C the medium was replaced with fresh medium and these cells were then treated as normal.

### 2.9.4 Determining cell number.

To aid the reproducibility of the transfection method (see 2.9.5) the number of cells per ml was determined by using a haemocytometer. Cells were detached from the flask according to the passage protocol (see 2.9.2) and irrespective of the total volume in which the cells had been resuspended, the number of cells $x10^4$ present in 1ml was represented by the number of cells counted under the microscope within the 5x5 square grid of the haemocytometer. The following equation was then used to calculate the volume of cell solution required to give a specific number of cells per well:

$$\text{Cells to be added } (\mu l) = \frac{\text{Specific cell density}}{\text{Cell number per ml}} \times 1000$$

To a 6 well microtitre plate, 3ml of antibiotic free DMEM media was added per well. Then cells were added at the relevant cell density, which was typically between 1-6 x $10^5$ cells per well. The cells were then incubated at 37°C overnight.

### 2.9.5 Transfecting cells using Lipofectamine™ 2000.

The transfection of tissue culture cells requires that the cells be seeded at a specific cell density 24 hours before transfection so that specific cell confluency can be achieved. The Lipofectamine 2000 (Invitrogen) transfection protocol advised that cells should be 90-95% confluent at the time of transfection. However, due to the very high growth rate of the mouse lens epithelial cell line, a lower confluency was favoured. Therefore, the cells were seeded so that they would be roughly 50% confluent at the time of transfection (1-2 x $10^5$ cells / well for a 6-well plate, seeded 24 hours before transfection – see 2.9.4). The optimised transfection protocol involved transfecting 2.5µg of DNA into cells using 6µl of Lipofectamine 2000. Both the 2.5µg of DNA and the 6µl of Lipofectamine 2000 were each diluted in 250µl Opti-MEM I Reduced Serum Media. The diluted Lipofectamine 2000 was incubated at room temperature for 5mins and then mixed with the diluted DNA, and incubated for a further 20mins at room temperature. The entire 500µl of DNA/Lipofectamine mix was then added to a well in a 6-well plate, which contained 3ml of fresh antibiotic free DMEM media. The cells were then left to incubate at 37°C for 48 hours, although the medium was replaced with fresh medium after 24 hours transfection.

## 2.10 Protein methods.

| | |
|---|---|
| **Acrylamide (Bio-Rad)** | 30% Stock Acrylamide/Bis Solution,<br>37.5:1 Acrylamide: *N,N'*-methylenebisacrylamide |
| **Blocking Buffer** | 5% w/v Fat free powdered milk (Marvel),<br>0.2% v/v Tween 20,<br>in PBS,<br>Make up fresh and store at 4°C |
| **Laemmli Running Buffer**<br>**(5x Stock)** | 125mM Tris,<br>1.25M Glycine,<br>0.5% (v/v) SDS,<br>Store at 4°C |
| **Laemmli Sample Buffer**<br>**(2x Stock)** | 10% (v/v) Glycerol,<br>4% (w/v) SDS,<br>60mM Tris, pH 6.8,<br>0.1% (w/v) Bromophenol Blue,<br>Add 200mM (dithiothreitol) DTT immediately before use. |
| **Ponceau S Stain** | 0.2% (w/v) Ponceau S,<br>in 3% (v/v) Glacial acetic acid. |
| **Running Gel (10%, 12%, 18%)** | 10%, 12% or 18% Acrylamide/Bis Solution,<br>375mM Tris, pH 8.8,<br>0.1% SDS,<br>0.1% (w/v) Ammonium persulphate (APS),<br>Typically made up to 10mls with $dH_2O$,<br>4µl TEMED.<br>Overlaid with $dH_2O$, allowed to set in plates (~30mins) |

| | |
|---|---|
| **Stacking Gel (5%)** | 5% Acrylamide/Bis Solution,<br>125mM Tris, pH 6.8,<br>0.1% SDS,<br>0.1% (w/v) APS,<br>Typically made up to 3mls with $dH_2O$,<br>3µl TEMED.<br>Comb inserted and allowed to set (~10mins). |
| **Transfer Buffer** | 25mM Tris,<br>200mM Glycine,<br>20% (v/v) Methanol,<br>stored at 4°C. |
| **Wash Buffer (TBS+ Tween)** | 10mM Tris-HCl, pH 7.5,<br>150mM NaCl,<br>0.2% (v/v) Tween 20,<br>in $dH_2O$,<br>Store at 4°C |

### 2.10.1 SDS-PAGE.

SDS-polyacrylamide gel electrophoresis was performed using the Bio-Rad Mini-PROTEAN® 3 cell. The Bio-Rad Mini-PROTEAN 3 cell instruction manual is comprehensive in its explanation of how to assemble and pour SDS-polyacrylamide gels and can be downloaded from the Bio-Rad website at http://www.biorad.com/cmc_upload/Literature/44432/4006157B.pdf. Separating and stacking gels were prepared as described in the instruction manual using the recipes given above. Although a 5% stacking gel was always used, a range of separating gels (10%-18%) was made to resolve protein bands of various sizes. High percentage separating gels were used to separate proteins with small molecular weights and vice versa. Protein samples were prepared by mixing protein lysate with an equal volume of 2x Laemmli sample buffer (containing 1mM DTT) and boiling at 100°C for 5mins. The reduced protein samples were then loaded into the wells of the stacking gel along with the correct markers. 2 types of markers were routinely used: Full Range Molecular Weight Rainbow Markers RPN800 (Amersham Pharmacia), 10µl commonly run on low percentage separating gels; and Low Molecular Weight Rainbow Markers RPN755 (Amersham Pharmacia), prepared in Laemmli sample buffer and then 10µl loaded on

high percentage separating gels. The gel was submerged in 1x Laemmli running buffer and electrophoresis performed at 150 volts for 1-2 hours. Progress of the migrating protein was monitored by the progress of the rainbow markers.

### 2.10.2 Electrophoretic Transfer to Hybond-P.

Electrophoretic transfer of protein was performed using the Bio-Rad Trans-Blot Semi-Dry Electrophoretic Transfer Cell. The Bio-Rad Trans-Blot SD cell instruction manual is comprehensive in its explanation of how to assemble a 'transfer sandwich' and can be downloaded from the Bio-Rad website at http://www.biorad.com/cmcupload/Literature/12490/M1703957B.pdf.

Following SDS-PAGE (see 2.10.1), the gel was equilibrated in transfer buffer for 15mins. An appropriately sized piece of polyvinylidene difluoride (PVDF) Hybond-P membrane (Amersham Pharmacia Biotech) was pre-wetted in methanol for 10secs, rinsed in distilled water and then was also incubated in transfer buffer for 15mins. A 'transfer sandwich' was then assembled on the Trans-Blot SD Transfer Cell, which was run at 15 volts for 1 hour. Following transfer it was ascertained whether protein had successfully been transferred by staining the Hybond-P membrane with Ponceau S. The membrane was stained with Ponceau S for 10secs and then the protein bands were visualised by washing 3 times in $dH_2O$. The membrane was then incubated overnight in blocking buffer at 4°C in preparation for western blotting (see 2.10.3).

### 2.10.3 Western Blotting.

Following overnight incubation in blocking buffer at 4°C, the membrane was then incubated with the appropriate primary and secondary antibodies (see Tables 2.9 and 2.10 respectively). The volume of blocking buffer that was used to dilute the primary antibody was 0.1ml/cm$^2$ membrane. The membrane was then sealed into a plastic pocket and incubated for 2 hours with end over end mixing at 4°C. The membrane was then removed from the plastic pocket and rinsed 3 times with wash buffer, followed by 1x 50ml wash for 15mins and 2x 50ml washes for 5mins each. Then the appropriate HRP-conjugated secondary antibody was diluted in 20ml of blocking buffer and incubated with the membrane at room temperature for 90mins. Finally the membrane was washed as before and then the signal was detected using ECLplus (Amersham Biosciences), according to the manufacturer's instructions. When ECL detection was used,

MACO-X-RAY HS90 medical film (Scientific Laboratory Supplies, SLS) was exposed and developed using a Hyper Processor (Amersham Life Science).

| Name. | Antibody. | Company. | Dilution. |
|---|---|---|---|
| C' PAX6 | Rabbit anti-PAX6 polyclonal antibody | Chemicon International | 1:1000 |
| GST | Goat anti-GST polyclonal antibody | Amersham | 1:4000 |
| FLAG | Mouse anti-FLAG® M2 Monoclonal antibody | Sigma | 1:1000 |
| PAX6 | Cocktail of 4 PAX6 antibodies: | | |
| | Chick PAX6 aa 1-223 antibody | DSHB* | 1:100 |
| | Human PAX6 aa 1-206 Sera AD 2.35.8 | Dieter Engelkamp | 1:50 |
| | Human PAX6 aa 1-206 Sera AD 2.3.7B | Dieter Engelkamp | 1:50 |
| | Human PAX6 aa 1-206 Sera AD 2.38.9B | Dieter Engelkamp | 1:50 |

**Table 2.9 Primary antibody dilutions.**

This table shows the primary antibodies that were used. The PAX6 cocktail is a mix of 4 different antibodies: An antibody from the Developmental Studies Hybridoma Bank (DHSB) designed against the first 223 amino acids of chick PAX6; 3 different monoclonal antibodies made by D. Engelkamp against the first 206 amino acids of human PAX6 (Engelkamp *et al.*, 1999).

| Name. | Antibody. | Company. | Dilution. |
|---|---|---|---|
| C' PAX6 | anti-rabbit HRP conjugated | DAKO | 1:5000 |
| GST | anti-goat HRP conjugated | DAKO | 1:10 000 |
| FLAG | anti-mouse HRP conjugated | Sigma | 1:80 000 |
| PAX6 | anti-mouse HRP conjugated | Sigma | 1:80 000 |

**Table 2.10 Secondary antibody dilutions.**

### 2.10.4 Large scale protein preparation from cells.

**Lysate Buffer**

150mM NaCl,
50mM Tris, pH 7.5,
1% NP-40 (Non-idet P),
0.5% Sodium Deoxycholate,
0.1% SDS
Store at 4°C,
Add 1x protease inhibitors before use.

Protein was isolated from the lens epithelium cell lines (see Table 2.8) so that they could be used as positive controls when Western Blotting with PAX6 antibodies. The method used to isolate protein was very similar to that routinely used as part of the immunoprecipitation protocol. A T75 flask of each cell type was grown until fully confluent (see 2.9.1), washed with 12ml V3 media and detached from the flask in the normal way using 12ml T:V media (see 2.9.2). 6ml of the dissociated cells were then combined with 24ml of DMEM full media and spun for 5mins at 750x*g*. The supernatant was then removed, the cell pellet resuspended in 250µl of Lysate Buffer and the mixture sonicated on ice for 5mins to reduce the viscosity of the cells. The sample was then boiled for 5mins and spun at 12 000x*g* for 5mins, and the supernatant was removed to a fresh 1.5ml eppendorf. The protein concentration was determined using the Bio-Rad DC protein assay (see 2.10.5) and the protein was stored in aliquots at -70°C.

### 2.10.5 Quantification of protein by use of the Bio-Rad DC Protein Assay.

Protein concentrations were determined using the Bio-Rad DC Protein Assay detection kit, as per the manufacturer's instructions. Briefly, 5µl sample was mixed with 25µl Reagent A, 200µl Reagent B and then incubated for 15mins. The absorbance at 750nm was read using a UV spectrophotometer. A BSA (bovine serum albumin) standard curve was generated each time the assay was performed by plotting absorbance values obtained from a range of BSA dilutions from 0-1.5mg/ml, and was used to determine the concentration of protein samples.

### 2.10.6 Subcellular Proteome Extraction Kit (Calbiochem®).

This kit was used to extract 4 subcellular fractions from an 80% confluent T25 flask of mouse lens epithelium cells (mv+, see Table 2.8). The 4 subcellular fraction contained: cytosolic proteins; membrane and organelle proteins; nucleic proteins; and cytoskeletal matrix proteins.

The subcellular proteome extraction kit protocol was followed exactly for adherent cells, as mv+ cells are adherent cells, and as recommended in the protocol when the cells began to detach after incubation with extraction buffer I the suspension cell protocol was followed from then on.

Samples were prepared for analysis by SDS-PAGE by combining 15µl of each subcellular fraction with an equal volume of 2x Laemmli sample buffer (containing 1mM DTT) and boiling at 100°C for 5mins. 10µl of each sample was then loaded onto a 12% SDS-polyacrylamide gel (see 2.10.1).

### 2.10.7 *In vitro* translation of $^{35}S$ radioactively labelled protein.

$^{35}S$ radioactively labelled protein was produced from QIAquick cleaned PCR amplified DNA and control vector DNA using the TnT® Coupled Reticulocyte Lysate System (Promega). *In vitro* translation reactions were initially made up to a volume of 23µl with nuclease free water and: 2µl of TnT reaction buffer; 1µl of T7 RNA polymerase; 1µl of amino acid mixture minus methionine (1mM); 1µl of RNasin ribonuclease inhibitor (40U/µl); and either 0.5µg of control vector DNA (0.5µg/µl) or approximately 200ng of sample PCR DNA (~20ng/µl). The reaction was then transferred to a fume hood and 2µl of $^{35}S$ methionine (20µCi of Amersham Pharmacia Redivue™ L-[$^{35}S$]methionine) was added to the reaction with adequate radioactive shielding. The total reaction volume was then made up to 50µl by adding 25µl (50% v/v) of TnT rabbit reticulocyte lysate with adequate radioactive shielding. The *in vitro* translation reaction was then incubated at 30°C for 90mins and then stored overnight at -20°C before being used in the Matchmaker co-IP kit (see 2.11.5).

## 2.11 Co-Immunoprecipitation.

| | |
|---|---|
| **Lysis Buffer (1% NP-40)** | 150mM NaCl,<br>50mM Tris, pH 8,<br>1% NP-40 (Non-idet P),<br>Store at 4°C, |
| **Complete Protease Inhibitor Cocktail (25x Stock)** | 1 tablet (Roche),<br>Dissolved in 2ml $dH_2O$.<br>Store at -20°C for up to 4 weeks. |

### 2.11.1 Cell Lysis.

After 48 hours transfection (see 2.9.5) the cells were washed in 1ml of room temperature PBS. The transfected cells were then treated with 250µl of chilled lysis buffer containing 1x protease inhibitors (PI), and incubated on ice at 4°C for 30mins, with occasional rocking. The cell lysate was then transferred to a chilled 1.5ml eppendorf and spun at 10 000x*g* for 10mins at 4°C. The supernatant was then removed to a fresh 1.5ml eppendorf. At this stage lysate was either used for pre-clearing and immunoprecipitation (see 2.11.3 and 2.11.4) or alternatively 15µl of lysate was added to 15µl of 2x Laemmli sample buffer (containing 1mM DTT), boiled for 5mins and then 10µl was analysed by SDS-PAGE and Western Blotting (see 2.10.1 to 2.10.3).

### 2.11.2 Preparing protein G.

Protein G sepharose fast flow beads were used in immunoprecipitation, as protein G is able to facilitate the purification of specifically tagged proteins. Protein G can bind to a certain antibody that recognises the tagged protein, and then this complex can be purified from a complex mixture of proteins using centrifugation and washing steps. The protein G sepharose was supplied in 20% ethanol and so needed to be washed 3 times in lysis buffer + PI before use.

### 2.11.3 Pre-clearing Lysate.

To reduce non-specific background bands, the cell lysate was first pre-cleared by incubating with protein G beads only. 200µl of the cell lysate was mixed with 50µg of protein G and made up to 500µl with lysis buffer + PI. A negative control that only contained lysis buffer + PI and

protein G was also always included. The pre-cleared lysates were then incubated at 4°C for 1 hour with end over end mixing.

### 2.11.4 Immunoprecipitation.

After the pre-clearing step the samples were spun at 10 000x*g* for 1min at 4°C. The supernatant was then poured into a fresh 1.5ml eppendorf and combined with 15μg of protein G and 0.1μg of goat anti-GST polyclonal antibody (Amersham). The samples were then incubated at 4°C for 2 hours with end over end mixing. It is vital at this stage that the experiment be kept chilled to reduce the amount of background bands. Therefore, the immunoprecipitation samples were spun at 10 000x*g* for 1min at 4°C and washed with 1ml of chilled lysis buffer + PI. The washing step was repeated 4 times so that the beads were thoroughly washed to remove any unbound protein. After the washing steps the protein G beads were suspended in 30μl of 2x Laemmli sample buffer (containing 1mM DTT) and boiled for 5mins to remove the bound protein from the protein G beads. The sample was then spun at 10 000x*g* and 10μl of the supernatant was analysed by SDS-PAGE and Western Blotting (see 2.10.1 to 2.10.3).

### 2.11.5 Matchmaker Co-IP Kit (BD Biosciences).

The Matchmaker co-immunoprecipitation kit was followed exactly as outlined in the BD Biosciences protocol. *In vitro* translated (see 2.10.7) protein was used to perform *in vitro* co-immunoprecipitation; however the specific details of proteins used in these reactions can be seen in Chapter 8.9. In brief $^{35}S$ radioactively labelled bait and library proteins were incubated together to allow interaction, and then either the cMyc or HA antibody was added to purify the proteins. Negative controls were also included to identify if there was non-specific binding of the incorrect antibody and epitope (i.e. cMyc purifying library protein containing the HA epitope). Protein A beads were then added to isolate the co-immunoprecipitation antibody complex and then a series of washes were performed. The washed beads were then combined with 2x Laemmli buffer, boiled and analysed by SDS-PAGE (see 2.10.1). After electrophoresis the gel was fixed and amplified and then dried at 80°C under constant vacuum. The gel was then exposed to X-ray film (same as in western blot, see 2.10.3) for up to 11 days at room temperature.

# Chapter 3 Defining the C terminal region of PAX6

## 3.1 Introduction.

As already discussed in the main introduction (see 1.76 and 1.8) it has been inferred from examining several mutations and from amino acid alignments that the C terminal portion of the PST domain seems to be very important for the normal function of PAX6. Therefore, the main goal of my PhD has been to investigate this C terminal end of PAX6.

## 3.2 BLAST searches.

At the start of my project the C terminal region of PAX6 was uncharacterised in terms of structure and function and so my first task was to define this C terminal region. This problem was addressed by performing BLASTP searches of the protein sequence database using the PST region of human PAX6 i.e. the amino acid sequence after the ASNT motif (highlighted in blue in Figure 3.1).

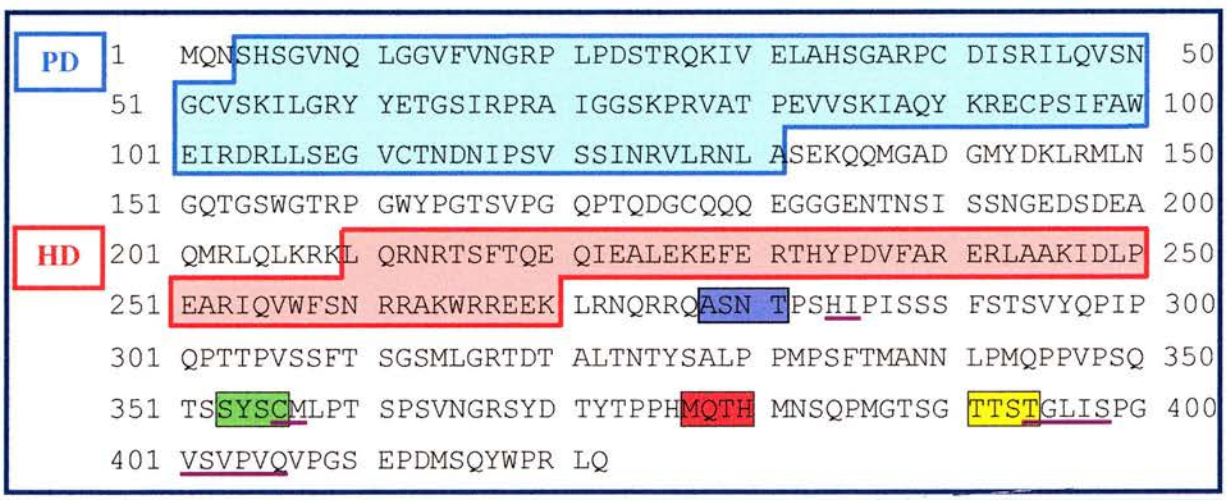



**Figure 3.1 The PAX6 protein and C terminal peptides.**

This shows the amino acid sequence of the PAX6 protein. The paired domain (PD) and the homeodomain (HD) are highlighted in blue and red boxes respectively. Also the first 4 amino acids of each of the 4 truncated C-terminal PAX6 peptides are marked. The blue ASNT peptide corresponds to the whole PST region; the green SYSC peptide corresponds to part of the PST region (containing a small beta sheet); the red MQTH peptide corresponds to a shorter portion of the PST region (without this beta sheet); and the yellow TTST peptide is part of the PST region that has been defined as the C terminal region of the PAX6 protein on the basis of secondary structure information and sequence alignments. Purple lines indicate β-sheets in the PST region.

The purpose of this search was to look for highly conserved amino acid motifs within the PAX6 PST domain of different organisms. Highly conserved regions of amino acid sequences often indicate areas of functional importance within a protein (Johnson and Lehtonen, 2000). The preliminary BLAST search showed that there was a conserved region within the last 42 amino acids of PAX6. This sequence was entered into another BLAST search to further investigate this region. This second search revealed that there was a very highly conserved region at the C terminus of PAX6 that was common between humans and very diverse organisms including axolotl (*Ambystoma mexicanum*), amphioxus (*Branchiostoma lanceolatum*) and common sea urchin (*Paracentrotus lividus*). To better illustrate this conserved region multiple protein sequences were aligned using the CLUSTAL program. This resulted in a clear indication of the most highly conserved region of the PAX6 C terminus – the "GLISP" motif (illustrated by the stars underneath the aligned sequences in Figure 3.2).

```
CLUSTAL W (1.74) multiple sequence alignment


cavefish2       MNSQSMATSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
zebrafish2      MNSQSMAASGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
chick2          MNSQPMGTSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
human2          MNSQPMGTSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
mouse2          MNSQPMGTSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
xenopus2        MNSQPMGTSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
axolotl2        MNSQSMGTAGATSTGLISPGVSVPVQVPGSEPDMS----QYWPRLQ
medaka2         MNSQSMTTSGTTSTGLISPGVSVPVQVPGSEPDMS----QYWSRLQ
urchin2         ------HMNSACNNGLISPGVSVPVQVPGSSPADTHPSHQYWPRIQ
amphioxus2      PGHSHMSHANGGSAGLISPGVSVPVQVPGAVTEEM-TSQPYWPRIQ
squid2          ---MTPHSTNGASTGLISPGVSVPIQVPGGGPQDM--STHHMSS--
                          .  . **********:****. .        : .
```

**Figure 3.2 Alignment of terminal 42 amino acids of human PAX6.**

This shows a CLUSTAL alignment of the terminal 42 amino acids of human PAX6 and PAX6 from other species. The stars underneath the alignment indicate invariant amino acids, whereas ':' indicates that the substitution is highly similar and '.' indicates that the substitution is only moderately similar.

Although this conserved GLISP peptide had been identified there was still some doubt as to where the cut off for the C terminal region should be. Therefore the amino acid sequence of the GLISP motif was put through a secondary structure predicting program called JPRED. This showed that this region contained a single predicted beta sheet (see Figure 3.3).

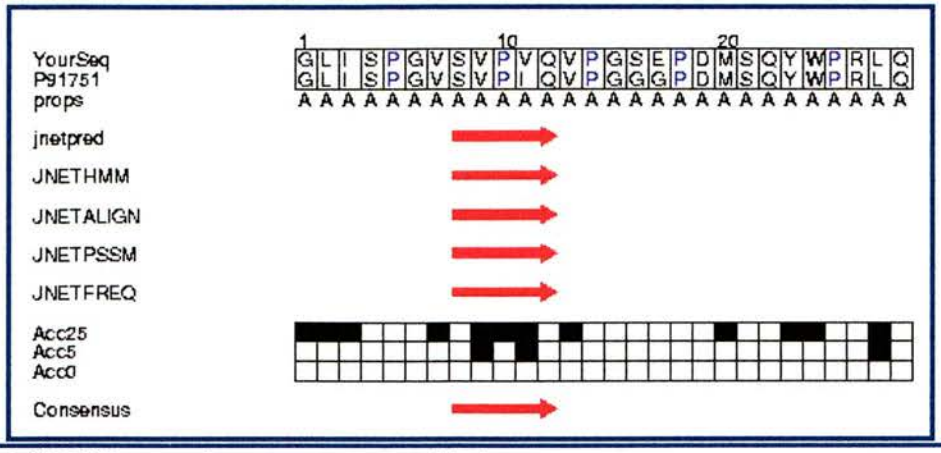



**Figure 3.3 GLISP secondary structure prediction.**

This shows the secondary structure as predicted by JPRED. The single beta sheet can be seen represented as the red arrow. The human PAX6 amino acid sequence is on the top line indicated by 'YourSeq.'

However, when the ASNT peptide was put through this program it was seen that there were actually 2 beta sheets at the GLISP region, one in the 'SVPVQ' peptide and one in the 'GLISP' peptide. This suggested that amino acids just N terminal of GLISP were required to seed the β-sheet. By gradually increasing the size of the GLISP peptide, in the N terminal direction, I found the exact cut-off that corresponded to this 2-beta sheet region at the C terminus of PAX6 (see Figure 3.4).

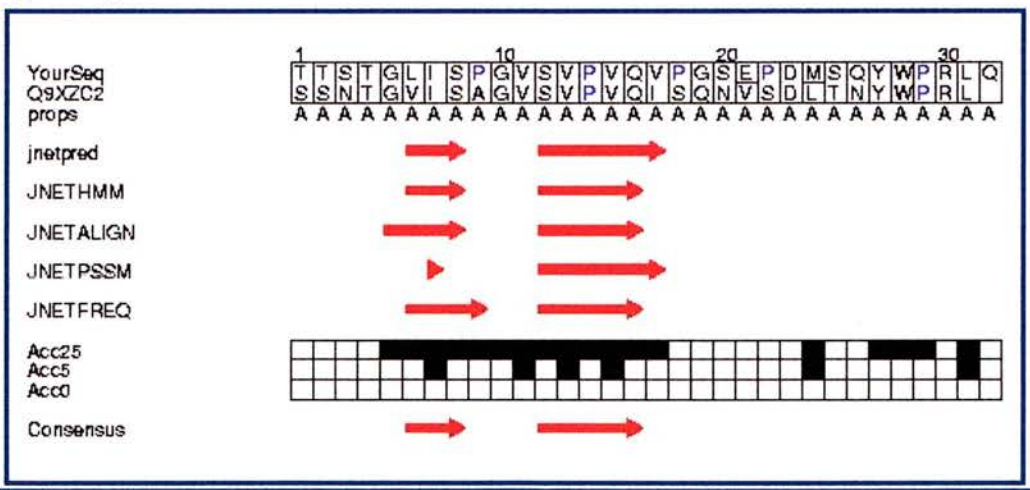



**Figure 3.4 TTST secondary structure prediction.**

This shows the secondary structure as predicted by JPRED. This shows that 2 beta sheets (represented by red arrows) are predicted when the sequence of PAX6 is extended 4 amino acids back from the conserved region. The human PAX6 amino acid sequence is on the top line indicated by 'YourSeq.'

Therefore on the basis of sequence conservation and secondary structure prediction, I have classed this TTST peptide as representing the C terminal domain of PAX6.

## 3.3 Conservation of the C terminal domain over evolution.

The discovery that the C terminal amino acids of the PST domain have been conserved over evolution was an interesting result, and the classification of the peptide starting with the amino acids TTST as this C terminal domain was valid when taking into consideration the results of the secondary structure prediction.

However, as an interesting aside I investigated the hypothesis that the C terminal region of PAX6 is important due to the retention of this specific secondary structure over evolution. To investigate this I took all of the sequences from the CLUSTAL alignment and put them through the JPRED secondary structure prediction program.

Unsurprisingly, the majority of the sequences had almost identical secondary structures predicted due to them only varying from human PAX6 by 1 or 2 amino acids. However, the sequence of amphioxus (*Branchiostoma lanceolatum*), california market squid (*Loligo opalescens*) and common sea urchin (*Paracentrotus lividus*), which are the most distantly related to the human PAX6 protein in evolutionary terms, are roughly 35% different within the region classified as the C terminal domain. However, it is interesting to see that even though the sequence is quite different the secondary structure is retained (see Figures 3.5 to 3.7).

I realise that in fact it is the highly conserved region of the C terminal domain that gives rise to the characteristic secondary structure and that this will therefore, be conserved over evolution. Even though the sequence surrounding this very highly conserved region has obviously altered during evolution, this has not altered the secondary structure. This retention of sequence and secondary structure conservation, therefore, gives even more strength to the explanation of a C terminal domain in PAX6. It is tantalising to think that the conservation of these amino acids is due to the C terminal domain having a specific function.

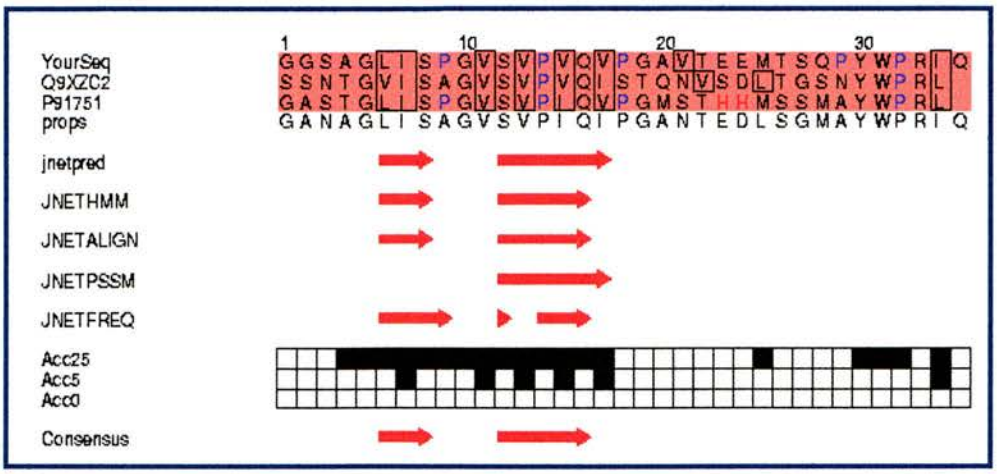



**Figure 3.5 Amphioxus secondary structure.**

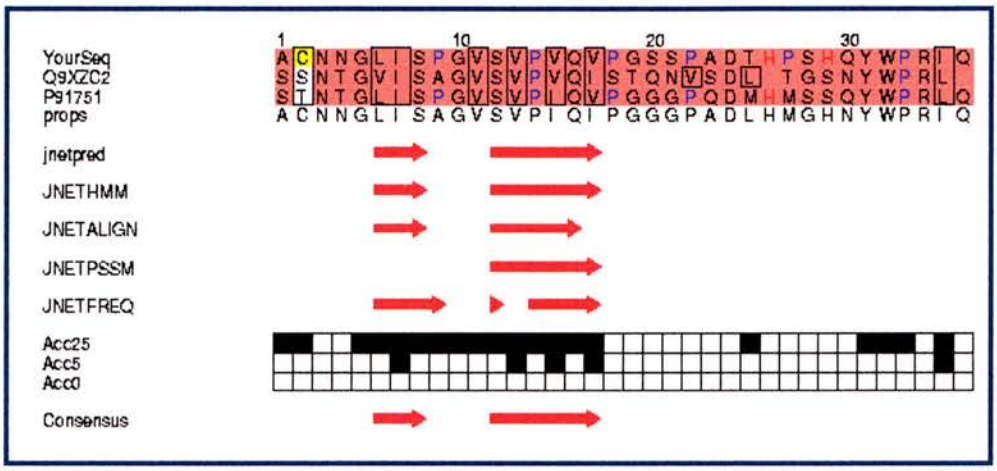



**Figure 3.6 Common sea urchin secondary structure.**

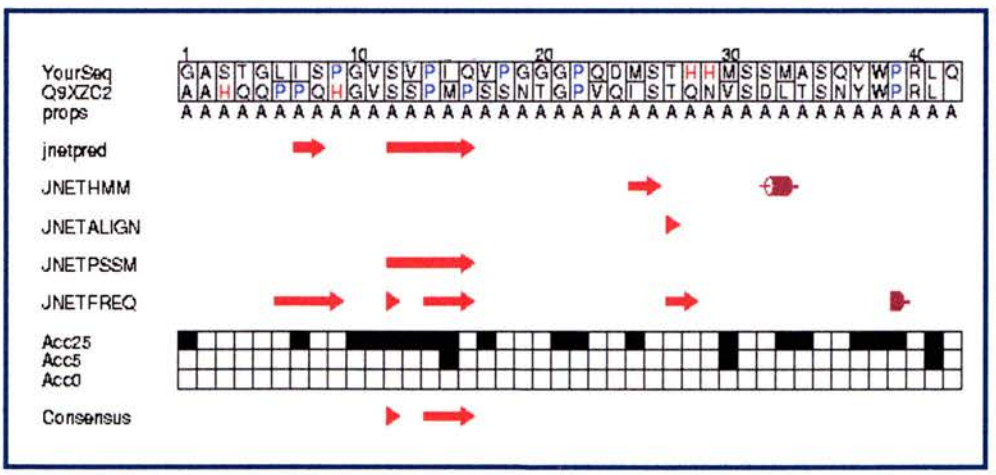



**Figure 3.7 California market squid secondary structure.**

# Chapter 4 Construct Cloning

## 4.1 Introduction.

Having defined the C terminal region of the PAX6 protein the next stage of my research was to make constructs that would be interesting to study. Although the specifics of each cloning strategy that I will outline are particularly varied there are many aspects that are very similar. Therefore, rather than repeating the same explanations for each cloning strategy, I will outline these commonalities first.

### 4.1.1 Common cloning strategies.

Subcloning steps in all of my cloning strategies were generated in one of two ways: either by ligation of compatible restriction digested DNA; or by generating specific DNA fragments by PCR. The DNA of all the clones produced by ligation was prepared using the Qiagen Miniprep kit and simply verified by diagnostic restriction digest. However, fragments that were generated by PCR were first cloned into pGEM-T (see 2.4.8) and insert-containing clones were identified by a β-Gal assay. The DNA of these clones was then prepared using the Qiagen Miniprep kit (see 2.2.1) and the presence and correct orientation of the PCR fragment was confirmed by diagnostic restriction digest. The clones were then verified by sequencing using T7 and SP6 sequencing primers (see Table 2.7) and Big Dye kit (see 2.7.1) to ensure that there were no PCR generated errors.

The DNA of all the verified clones was prepared using the Qiagen Midiprep kit (see 2.2.2) and glycerol stocks (see 2.6.5) were made for each subcloning stage product as well as the final complete clone.

## 4.2 Cloning truncated *PAX6* into pDBLeu.

I decided to make 4 *PAX6* constructs in the pDBLeu expression vector. This vector is used to clone a cDNA sequence of interest in frame with the GAL4 DNA binding domain for use in the yeast 2-hybrid system. Constructs that correspond to the PST region of PAX6 and the newly defined C terminal domain were made and used to screen a cDNA library using the yeast 2-hybrid system (see Chapter 5).

I also decided to prepare 2 other clones containing the green SYSC and the red MQTH peptides (see Figure 3.1). These were to be used in pairwise interactions with the isolated interactors from

the yeast 2-hybrid screen. These truncated peptides were chosen due to a combination of patches of conserved sequence seen around the SYSC and MQTH regions of the sequence in the BLAST search against the PST region and secondary structure information. The SYSC peptide was chosen as it contains the 3' end of the PST domain, which contains a small predicted beta sheet. The MQTH peptide is even further towards the 3'end of the PST region but does not contain this small beta sheet (see Figure 3.1)

Having decided on these initial *PAX6* clones, I then cloned the corresponding *PAX6* cDNA fragments into the pDBLeu expression vector using the following cloning strategy.

PCR primers were designed to produce the relevant truncated *PAX6* cDNA fragments with a *Sal*I restriction site at the 5' end and a *Not*I site at the 3' end (see Table 2.3). The correctly sized PCR fragments were cloned into pGEM-T and, following sequence verification, were cut with *Sal*I and *Not*I and sub-cloned into pDBLeu.

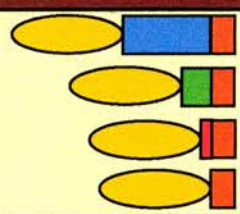
| Construct Produced. | Description. | Schematic Diagrams. |
|---|---|---|
| pDBLeu_ASNT_normal | PAX6 PST domain only. |  |
| pDBLeu_SYSC_normal | Patches of conserved sequence. | |
| pDBLeu_MQTH_normal | Patches of conserved sequence. | |
| pDBLeu_TTST_normal | PAX6 C terminal domain only. | |

**Table 4.1 pDBLeu constructs.**

This table shows the 4 different truncated proteins that were cloned into pDBLeu. The pDBLeu vector produces a protein coupled to the GAL4 DNA Binding Domain, which is illustrated in the schematic diagram. The GAL4 DNA binding domain is shown as a yellow oval. The C terminal domain is shown as an orange rectangle. Blue, green and red rectangles represent the PST domain, the SYSC peptide and the MQTH peptide respectively.

## 4.3 Cloning full length *PAX6* into pDBLeu.

In addition to these truncated proteins I also cloned full length *PAX6* into the pDBLeu expression vectors as a control for pairwise comparisons (see Chapter 6).

Full length *PAX6* was cloned into pDBLeu by a combination of subcloning and PCR. Specifically, the majority of full length *PAX6* was cut out of a construct containing the full length *PAX6* cDNA (-5a isoform) in pBluescript® SK+/- (clone EE4, no.53, a gift from Dieter Engelkamp), using *Nhe*I and *Not*I double digest and ligated into pDBLeu cut with the same restriction enzymes. The most 5' portion of *PAX6* was created by PCR, using EE4 as template DNA, producing a fragment that extends from an artificially inserted 5' *Sal*I site up to the *Nhe*I site within the *PAX6* paired domain. The sequence verified PCR clone was then inserted into the previously created clone using a *Sal*I and *Nhe*I double digest. The missing 300bp of *PAX6* was then cut out of the EE4 clone, by cutting with *Nhe*I, and then inserted into the intermediate clone to generate the final confirmed pDBLeu full length *PAX6* clone.

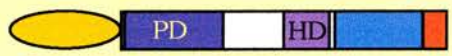
| Construct Produced. | Description. | Schematic Diagram. |
|---|---|---|
| pDBLeu_PAX6_Full | Corresponds to the entire –5a *PAX6* isoform. |   |

**Table 4.2 Full length PAX6/pDBLeu construct.**

This table shows the full length *PAX6* protein that was cloned into pDBLeu. The schematic diagram illustrates the GAL4/PAX6 fusion protein showing its various domains. PD, paired domain; HD, homeodomain. Other domains as in Table 4.1.

## 4.4 Cloning the Q422R PAX6 mutant into pDBLeu.

Substitution of the C terminal amino acid (Q422R) of PAX6 causes eye malformations including aniridia and uveal ectropion (Azuma and Yamada, 1998; Singh *et al.*, 2001). The Q422R mutation was incorporated into the truncated pDBLeu constructs (see Table 4.1) to investigate any differences between normal and mutated PAX6. The cloning procedure to create these constructs was exactly the same as for the normal constructs (see 4.2); however a different 3' PCR primer was used to incorporate the A to G nucleotide change at position 1682 (see Table 2.3).

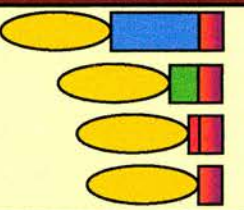
| Construct Produced. | Description. | Schematic Diagrams. |
|---|---|---|
| pDBLeu_ASNT_mutant | PST domain with Q422R substitution. |  |
| pDBLeu_SYSC_mutant | Containing Q422R substitution. | |
| pDBLeu_MQTH_mutant | Containing Q422R substitution. | |
| pDBLeu_TTST_mutant | C terminal domain with Q422R substitution. | |

**Table 4.3 Q422R mutant pDBLeu constructs.**

This table shows the 4 different Q422R mutant truncated proteins that were cloned into pDBLeu. The different shading of the C terminal domain in the schematic diagram (compared to the normal shading in Table 4.1) is meant to illustrate the single amino acid change at the very C terminus of the PAX6 protein. Other domains are as in Table 4.1.

## 4.5 Cloning WHEDA and REMAR mutant truncated *PAX6* into pDBLeu.

2 interesting C terminal mutations have been identified in aniridia patients called WHEDA and REMAR. WHEDA has the mutation 1615del10 (Heyman *et al.*, 1999) and REMAR has the mutation X423L (Sisodiya *et al.*, 2001). Both these mutations alter the reading frame of *PAX6* at the C terminus, causing read through of the normal stop codon into the 3' untranslated region of *PAX6*. The details of these mutations have already been discussed in the introduction (see 1.7.6).

pDBLeu constructs corresponding to the PST domain (ASNT) and C terminal domain (TTST) were prepared from the mutant and normal alleles of WHEDA and REMAR. The nested RT-PCR method (see 2.4.5) was used to generate all the TTST mutant and normal fragments and the ASNT WHEDA mutant fragments using the following second set of PCR primers: The *Sal*I conjugated 5' primer used to amplify the TTST truncated protein was also used for the TTST mutants (ST004, see Table 2.3); A 5' primer that corresponds to the *Nde*I site in *PAX6* was used for the ASNT mutant (ST015, Table 2.3); and a 3' primer, containing a *Not*I site, that was designed about 350bp downstream of the stop codon of *PAX6* was used for both TTST and ASNT mutants (ST013 for TTST and ST031 for ASNT, Table 2.3). Poly (A) addition (see 2.4.7) was carried out on these PCR fragments, which were then processed as normal. The mutant and normal fragments were identified by sequence analysis and then the verified PCR fragments were cloned into pDBLeu using a *Sal*I and *Not*I digest for all the TTST mutant and normal constructs and a *Nde*I and *Not*I digest for the ASNT WHEDA mutant construct.

The remaining ASNT clones were generated by the following method: the wild type REMAR and WHEDA constructs and mutant REMAR constructs were created by transferring a *Mlu*NI and *Not*I cut fragment from the TTST clones into pDBLeu_ASNT_normal (see Table 4.1) also cut with *Mlu*NI and *Not*I.

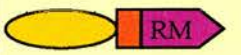
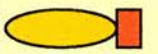
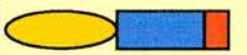
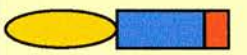
| Construct Produced. | Description | Schematic diagrams. |
|---|---|---|
| pDBLeu_TTST_WHmut | C ter domain with WHEDA mutant allele 1615del10. |   |
| pDBLeu_ASNT_WHmut | PST domain with WHEDA mutant allele 1615del10. |   |
| pDBLeu_TTST_RMmut | C ter domain with REMAR mutant allele X423L. |   |
| pDBLeu_ASNT_RMmut | PST domain with REMAR mutant allele X423L. |   |
| pDBLeu_TTST_WHwt | C ter domain with WHEDA wild type allele. |  |
| pDBLeu_ASNT_WHwt | PST domain with WHEDA wild type allele. |  |
| pDBLeu_TTST_RMwt | C ter domain with REMAR wild type allele. |  |
| pDBLeu_ASNT_RMwt | PST domain with REMAR wild type allele. |  |

**Table 4.4 WHEDA and REMAR mutant pDBLeu constructs.**

This table shows pDBLeu constructs corresponding to the PST domain (ASNT) and C terminal domain (TTST) that were prepared from the mutant and normal alleles of WHEDA (WH) and REMAR (RM). The arrow boxes in the schematic diagrams of these constructs illustrate the predicted elongated proteins due to the read through of the normal *PAX6* stop codon. The schematic diagrams for the wild type constructs are exactly as the corresponding normal PAX6 constructs in Table 4.1, as there will be no translation past the normal stop codon even though the 3' UTR sequence is included.

## 4.6 Cloning a minus C terminal domain PAX6 construct.

This method was used to make a pDBLeu_ASNT_normal construct (equivalent to the PST domain of PAX6) that was lacking the C terminal domain of PAX6. This construct was created as a control to investigate any differences in pairwise interactions with the isolated interacting proteins. It was hypothesised that this construct would cement the importance of the C terminal domain in the functioning of PAX6.

This construct was created by ligating a synthetic linker into pDBLeu_ASNT_normal construct, which had already been made (see Table 4.1). The annealed oligos (see 2.5.2) were diluted 1:100 and 4µl of this (~60ng linker) used to ligate into 35ng of *Nde*I, *Not*I cut pDBLeu_ASNT_normal (see 2.2.10) and transformed into XL1-Blue Subcloning-Grade Competent Cells (see 2.6.1).

| Construct Produced. | Description. | Schematic Diagram. |
|---|---|---|
| pDBLeu_ASNT-TTST | PST domain without the C terminal domain. |  |

**Table 4.5 Minus C terminal domain PAX6 construct.**

This table shows the minus C terminal domain PAX6 construct and this is illustrated by a schematic diagram. Domains as in Table 4.1.

# Chapter 5 Yeast 2-hybrid screening

## 5.1 Introduction.

Having successfully completed the cloning of the PAX6 PST (pDBLeu_ASNT_norm) and C terminal domain (pDBLeu_TTST_norm) constructs (see Table 4.1), the next step was to use these pDBLeu constructs to investigate interactions with other proteins. This was done by using the yeast 2-hybrid system to screen a mouse brain cDNA library (see 2.8.1).

### 5.1.1 Preparing the mouse brain cDNA library.

The cDNA library was prepared as described in the materials and methods section (2.8.1). The concentration of the resulting double stranded plasmid DNA was then quantified using spectrophotometric measurements (2.2.3) and the DNA diluted and stored in 1μg/μl aliquots. As a further control a *Not*I and *Sal*I double digest was performed (2.2.4) on 1μg of the cDNA library DNA to release the cDNA inserts from the vector. Therefore, if the correct DNA has been prepared then a pPC86 vector band of 7kb as well as a smear of DNA (which represents the variety of sizes of DNA fragments in the cDNA library) will be seen, as is shown in Figure 5.1.

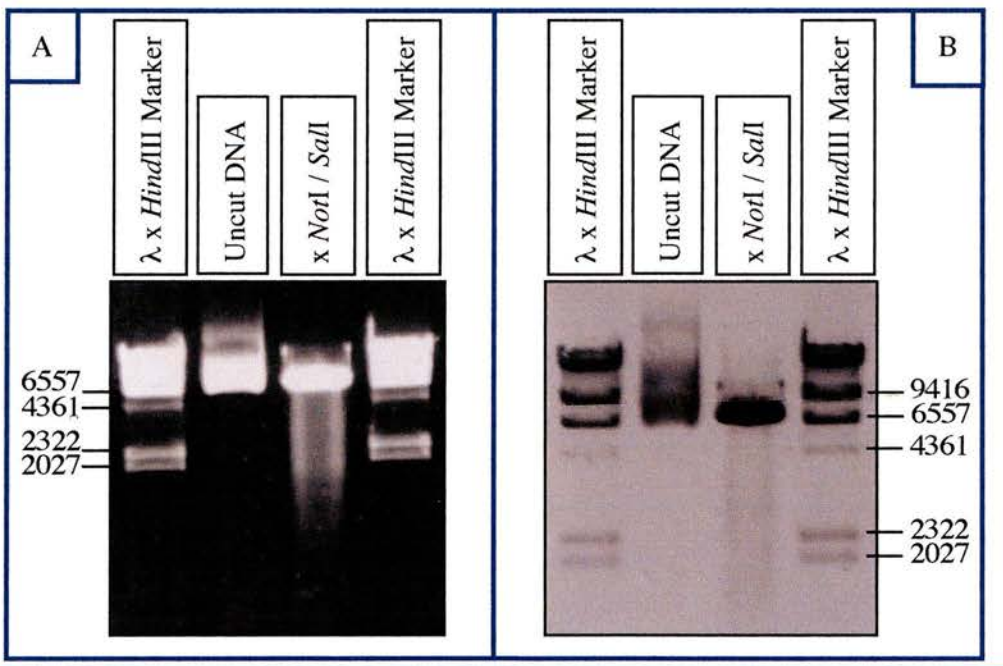



**Figure 5.1 Mouse brain cDNA library confirmation.**

The cDNA library DNA (1μg) was cut with *Not*I and *Sal*I and run out on a 0.8% agarose gel. 500ng of uncut cDNA library DNA and λ x *Hind*III Marker were also run out. A, This clearly shows the smear of the cDNA library fragments in the double digest lane. B, This gel has been run slightly further and clearly shows the 7kb pPC86 vector band in the double digest lane.

### 5.1.2 The yeast 2-hybrid system.

The yeast 2-hybrid system (see 2.8) utilises the DNA binding transcriptional activator, GAL4. The pDBLeu constructs that have been made produce truncated PAX6 proteins that are coupled to a GAL4 DNA binding domain (see Table 4.1), whereas the proteins produced by the cDNA library will be coupled to a GAL4 activation domain. Therefore, in the yeast 2-hybrid system, the GAL4-PAX6 fusion protein expressed from the pDBLeu construct will bind to the GAL4 binding site upstream of a reporter gene. However, unless there is interaction between the GAL4-PAX6 fusion protein and a protein from the cDNA library, there will be no expression of the reporter gene because a functional transcription factor will not be formed. This is outlined in the diagram below (see Figure 5.2).

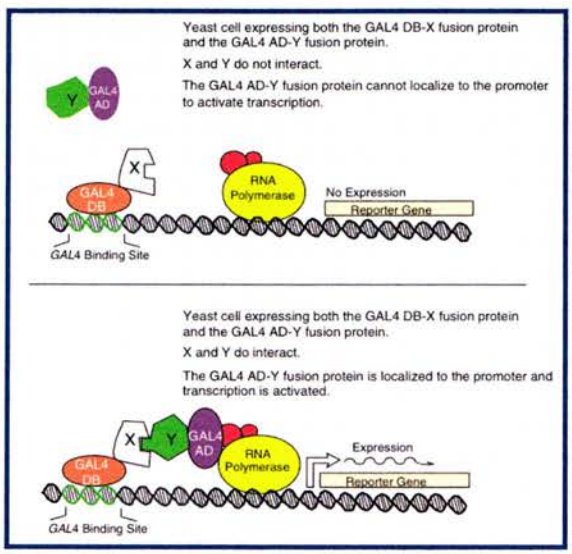



**Figure 5.2 Illustration of the 2-hybrid system.**

This shows how the 2-hybrid system works. My fusion proteins (X) are attached to a GAL4 DNA binding domain (GAL4 DB) and the cDNA library proteins (Y) are attached to a GAL4 activation domain (GAL4 AD). Therefore, there will only be reporter gene expression if proteins X and Y interact (*from ProQuest 2-hybrid instruction manual*).

## 5.2 Yeast 2-hybrid library screen.

In the yeast 2-hybrid system the first reporter gene that is utilised to screen the cDNA library is *HIS3*. However, before the mouse brain cDNA library was screened, the pDBLeu constructs were optimised for use with this reporter gene. As described in the materials and methods (see 2.8.4) a 3AT screen was performed on the pDBLeu constructs to eliminate any background self-activation of the *HIS3* reporter. This resulted in a 3AT concentration of 7.5mM for the pDBLeu_TTST_norm construct and 10mM for the pDBLeu_ASNT_norm construct required to eliminate basal levels of *HIS3* reporter gene expression.

The yeast 2-hybrid system was then used to screen the mouse brain cDNA library with the pDBLeu_TTST_norm and pDBLeu_ASNT_norm constructs, which correspond to the C terminal and PST domains of PAX6 respectively (see Table 4.1).

The cDNA library was screened 3 times using the C terminal domain construct (reasoning behind this discussed in 5.2.2) and screened once using the PST domain construct. For each library screen, it was estimated that $1x10^7$ to $1x10^8$ yeast 2-hybrid library clones were screened. In the C terminal domain screens it was recorded that small numbers (10-50) of *HIS3* positive colonies grew within 3 days incubation at 30°C after replica cleaning and the position of these colonies were marked. However, when the plates were incubated for longer thousands of *HIS3* positive colonies began to develop. In the PST domain screen the case was even more severe. Even before replica cleaning, 60 *HIS3* positive colonies had developed, whose position was marked, but on further incubation after replica cleaning although these 60 colonies re-grew, thousands of new *HIS3* positive colonies developed.

I could not discount the thousands of *HIS3* positive colonies but likewise could not spend endless amounts of time screening them all using the full screen. Therefore, I decided that it would be pertinent to randomly pick 475 of the thousands of *HIS3* positive colonies and test them using an X-gal assay (2.8.8). This would utilise the *lacZ* reporter gene and allow quick identification of *lacZ* and *HIS3* positive colonies due to the colour change (white to blue) that is central to the X-gal assay (see Figure 5.3).

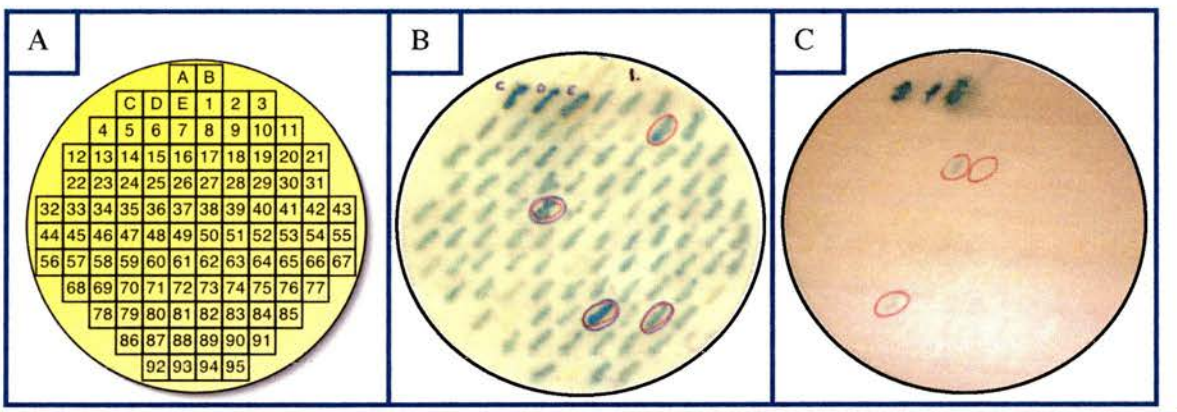



**Figure 5.3 High throughput screening of *HIS3* positive colonies.**

A, This is the template used to screen large numbers of *HIS3* positive colonies using an X-gal assay. The first 5 positions were the 5 yeast controls (A-E) followed by 95 *HIS3* positive colonies. B, This is one of the plates used to screen *HIS3* positive colonies from the PST domain library screen. The transactivation ability of the whole PST domain is clearly recognisable as the *lacZ* reporter gene is constitutively expressed. C, This shows screening of *HIS3* positive colonies from the C terminal domain library screen, which does not show constitutive transactivation of the *lacZ* reporter gene. In both B and C, colonies that first developed the positive blue colour in the X-gal assay are ringed.

The results of these high throughput X-gal assays were very interesting. Not only did they confirm that the majority of these later developing colonies were false positives but also with the *HIS3* positive colonies from the PST domain screen it was seen that all of the colonies developed a faint blue colour, suggesting that the PST fusion protein had innate transactivation activity. This is immediately noticeable when comparing the X-gal assay on *HIS3* positive colonies from the PST domain (B, Figure 5.3) and C terminal domain (C, Figure 5.3) screens as the colonies from the PST domain almost all have a positive blue colour. This transactivation ability of the PST domain, whilst already having been described (Tang *et al.*, 1998) and tested by other authors (Singh *et al.*, 1998), was investigated further and my results can be seen in Chapter 7.

Although the main function of this high throughput technique was to confirm that the majority of the later developing colonies were false positives, 2% of the tested *HIS3* positive colonies from the C terminal domain screen and 12% of those from the PST domain screen showed *lacZ* reporter gene expression in the X-gal assay. The *HIS3* positive colonies that first developed the positive blue colour in the X-gal assay (circled in Figure 5.3 panels B and C) were taken on to be screened using all 3 reporter genes. The amount of *HIS3* positive colonies, from the library screens and *lacZ* and *HIS3* positive colonies from the high throughput X-gal test plates that were taken on for further screening using the 3 reporter genes are outlined in Table 5.1.

| Library Screen. | *HIS3* positive. | *LacZ* and *HIS3* positive. |
|---|---|---|
| **PST domain** | 60 | 56 |
| **CT domain** | 95 | 9 |

**Table 5.1 Number of colonies investigated further using all 3 reporter genes.**

This table outlines how many positive colonies were taken on for further screening using all 3 reporter genes. There are 2 categories; *HIS3* positive, those *HIS3* positive colonies that grew on the screen plates before replica cleaning in the PST domain screen and within the first 3 days after replica cleaning in the C terminal (CT) domain screen; *LacZ* and *HIS3* positive, the *lacZ* and *HIS3* positive colonies that were identified by the high throughput X-gal assay.

### 5.2.1 Further investigation using all 3 reporter genes.

As outlined in the materials and methods section (2.8.7) further characterisation of the *HIS3* positive colonies was carried out using the 3 reporter genes *HIS3*, *URA3* and *lacZ* that are integrated into the yeast genome in this system. This allows 4 tests to be carried out on each

positive colony: 2 are simple replacement of essential amino acids (histidine and uracil) due to *HIS3* and *URA3* reporter gene expression, allowing growth of positive colonies on plates deficient of these amino acids; the X-gal assay which easily identifies positive colonies as they activate *lacZ* reporter gene expression and initiate a colorimetric change (white to blue); a test for weak interactions using repression of growth of yeast due to *URA3* reporter gene activation by positive colonies which converts 5FOA into toxic 5-fluorouracil.

Single colonies of the *HIS3* and *HIS3* and *lacZ* positive colonies from the main library screens were first obtained from glycerol stocks and master plates were set up (see 2.8.7). As illustrated in the materials and methods section, each master plate contained: 5 positive colonies that were to be tested plated in quadruplicate; a negative control that tests for self activation of reporter gene expression; and a full range of control yeast strains (A to E) which show increased activation of the various reporter genes.

The results obtained from this screening using all 3 reporter genes were then compared to the expected results as outlined in Table 5.2.

| -His | lacZ | -Ura | 0.2% 5FOA | Easily-Interpreted Phenotypes |
|---|---|---|---|---|
| no growth | | | | False positive/background |
| growth | blue | no growth | no growth | Interactor, probably weak |
| growth | blue | growth | no growth | Interactor |
| growth | white | no growth | growth | Non-interactor |

| -His | lacZ | -Ura | 0.2% 5FOA | One Inconsistent Phenotype |
|---|---|---|---|---|
| growth | blue | growth | growth | Probable interactor<br>Possible mixed population* |
| growth | white | growth | no growth | Probable interactor (check *lacZ* expression using CPRG assay) |

| -His | lacZ | -Ura | 0.2% 5FOA | Two Inconsistent Phenotypes |
|---|---|---|---|---|
| growth | blue | no growth | growth | Possible non-interactor<br>Possible weak interactor (look carefully for any inhibition on 5FOA)<br>Possible mixed population* |
| growth | white | growth | growth | Possible interactor (check *lacZ* expression using CPRG assay)<br>Possible mixed population*; look carefully for inhibition on 5FOA |
| growth | white | no growth | no growth | Probable weak interactor (check *lacZ* expression using CPRG assay; confirm 5FOA phenotype) |

*Mixed populations can result from numerous causes. Examples include: More than one AD-Y in the cell resulting in one population that contains AD-Y interacting with DB-X (no growth on 5FOA), and another population that contains AD-Y that does not interact with DB-X, giving rise to growth on 5FOA; mutations or instability in DB-X, AD-Y or *URA*3 reporter or promoter; carry over of cells from the -His+3AT plate that do not contain interacting DB-X:AD-Y.

**Table 5.2 Possible phenotypes from full screen.**

This table shows the entire spectrum of phenotypes that are possible from screening the positive colonies with the 3 reporter genes. *Taken from ProQuest 2-Hybrid system manual (Gibco-BRL).*

### 5.2.2 C terminal domain full screen results.

The following table (Table 5.3) summarises the information obtained from screening the 95 *HIS3* positive colonies and the 9 *HIS3* and *lacZ* positive colonies isolated from the C terminal domain library screen, using the full range of 3 reporter genes. The table is laid out to represent the table from the ProQuest 2-hybrid system manual (see Table 5.2).

| H | L | U | 5F | Phenotype | CT | CR | Seq |
|---|---|---|---|---|---|---|---|
| NG | | | | False positive | 16 | - | - |
| G | B | NG | NG | Interactor, probably weak | 12 | 8 | 4 |
| G | B | G | NG | Interactor | 2 | 1 | 1 |
| G | W | NG | G | Non-interactor | 50 | - | - |
| G | B | G | G | Probable interactor or mixed population | 2 | - | 2 |
| G | W | G | NG | Probable interactor | 4 | 1 | 3 |
| G | B | NG | G | Possible non-interactor, weak inter or mixed pop. | 5 | - | - |
| G | W | G | G | Possible interactor or mixed population | 6 | - | 4 |
| G | W | NG | NG | Probable weak interactor | 7 | 6 | 1 |

**Table 5.3 Full screen of C terminal domain positive colonies.**

This table shows the combined full screen results of all 3 of the C terminal domain library screens. The coloured boxes and phenotype descriptions refer to Table 5.2. G, growth identified as red; NG, no growth identified as green; B and W, blue and white colour in X-gal assay. CT, the number of positive colonies that correspond to each phenotype. CR, Some colonies could not be recovered from the first library screen (see main text). Seq, this shows the number of colonies in which the cDNA insert was sequenced. H, *HIS3*; L, *lacZ*; U, *URA3*; 5F, 0.2%5FOA (see 2.8.7).

I have presented the data from the C terminal domain full screens in combined form but there are certain discrepancies that need to be explained. When performing the first C terminal domain library screen, glycerol stocks were not made of the *HIS3* positive colonies and therefore some possibly interesting positive colonies were lost (hence Couldn't Recover column in Table 5.3). Therefore, this is the reasoning for performing another C terminal domain library screen. I screened the mouse brain cDNA library a further two times to recover more positive colonies

with phenotypes corresponding to interactors. As is demonstrated in Table 5.3 the only positive colonies that I did not mark for sequencing were those whose phenotype corresponded to non-interactors.

### 5.2.3 PST domain full screen results.

The following table (Table 5.4) summarises the information obtained from screening the 60 *HIS3* positive colonies and the 56 *HIS3* and *lacZ* positive colonies isolated from the PST domain library screen, using the full range of 3 reporter genes. The table is laid out to represent the table from the ProQuest 2-hybrid system manual (see Table 5.2).

| H | L | U | 5F | Phenotype | PST | HTX | Seq |
|---|---|---|---|---|---|---|---|
| NG | | | | False positive | - | - | - |
| G | B | NG | NG | Interactor, probably weak | 1 | 3 | 4 |
| G | B | G | NG | Interactor | 32 | 22 | 52 |
| G | W | NG | G | Non-interactor | 4 | - | - |
| | | | | | | | |
| G | B | G | G | Probable interactor or possible mixed pop. | - | 1 | - |
| G | W | G | NG | Probable interactor | 5 | 31 | 5 |
| | | | | | | | |
| G | B | NG | G | Poss. non-interactor, weak inter or mixed pop. | - | 1 | - |
| G | W | G | G | Poss. interactor or mixed population | 1 | 10 | 1 |
| G | W | NG | NG | Probable weak interactor | - | 5 | - |

**Table 5.4 Full screen of PST domain positive colonies.**

This table shows the results of the PST domain library screen. The coloured boxes and phenotype descriptions refer to Table 5.2. G, growth identified as red; NG, no growth identified as green; B and W, blue and white colour in X-gal assay. PST, the number of *HIS3* positive colonies that correspond to each phenotype. HTX, the number of *HIS3* and *lacZ* positive colonies that correspond to each phenotype. Seq, this shows the number of colonies' DNA that was sequenced. H, *HIS3*; L, *lacZ*; U, *URA3*; 5F, 0.2%5FOA (see 2.8.7).

As previously described for the C terminal library screen, by combining data from the *HIS3* positive colonies and the *HIS3* and *lacZ* positive colonies, certain discrepancies occur that need

to be further explained. In the PST domain full screen results the disparity occurs when considering which positive colonies I marked for sequencing. Therefore, I have separated the results of these full screens into 2 groups, the *HIS3* positive colonies identified by the main PST domain library screen and the *HIS3* and *lacZ* positive colonies that were identified from the high throughput X gal assay (see Table 5.4). In this way it can be clearly seen that *HIS3* positive colonies that correspond to interactors or probable interactors were marked for sequencing, although I was unable to isolate the DNA of 2 interactors. However, due to the large number of samples that were already marked for sequencing when the results of the *HIS3* and *lacZ* positive colonies were obtained, only the unequivocal interactors and weak interactors were marked for sequencing.

## 5.3 Summary of fully characterised positive colonies.

The mouse brain cDNA library was screened 3 times using the C terminal domain construct (pDBLeu_TTST_norm) and once with the PST domain construct (pDBLeu_ASNT_norm). The detailed results of the individual screens can be seen above, however, a more condensed version is shown in the table below (see Table 5.5).

| Library Screen. | Colonies to be fully screened | Number sequenced |
|---|---|---|
| **PST domain**  | 116 | 62 |
| **CT domain**  | 95 | 15 |

**Table 5.5 Condensed yeast 2-hybrid screening results.**

This table summarises the results of the yeast 2-hybrid library screens. For both the PST domain and C terminal (CT) domain library screens this table shows: the number of positive colonies that were further characterised using the full screen (utilising all 3 reporter genes); and the number of fully screened positive colonies' DNA that was taken on to be sequenced.

In total 77 positive colonies were identified as being interesting to investigate further, 62 from the whole PST domain screen and 15 from the C terminal domain screen. The positive colonies were re-grown and cDNA library plasmid DNA was made from the yeast cultures as outlined in the materials and methods section 2.8.9.

To try and ascertain the identity of the positive colonies a simple *Not*I and *Sal*I double digest was performed on isolated DNA of the 77 positive colonies. This produced a range of DNA

sized bands, which were useful once the colonies had been sequenced to check whether the nucleotide sequence that was pulled out of the database was roughly the same length as that of the cDNA fragment in the library. Obviously this is not foolproof as the cDNA in the library could be a fraction of the clone in the database or vice versa.

The resultant plasmid DNA was then sequenced using the specific sequencing primers for the pPC86 vector (see Table 2.7). The resulting DNA sequence traces were then checked and edited so that the pPC86 vector sequence was removed and only unambiguous nucleotide sequence remained. The trimmed DNA sequence was then used to search nucleotide databases using the BLASTN program (see 2.1.1). This revealed the following information.

## 5.4 Isolated Interactors.

The DNA sequence of 72 out of the 77 positive colonies identified entries in the nucleotide databases when searched using BLASTN. It was found that: the nucleotide sequence of 7 colonies aligned to various RIKEN cDNA clones of unknown identity from thymus, pancreas, lung, small intestine, brain and 2 from whole embryo; the sequence of one colony aligned to the mouse mitochondrial genome and the sequence of another colony aligned to the GAL4 vector.

However, the majority of the nucleotide sequences of the positive colonies aligned to nucleotide sequences that encode proteins: the sequence of 3 colonies aligned to ribosomal RNA genes; the sequence of 7 colonies aligned to nucleotide sequences that encode hypothetical proteins; 20 sequences each aligned to sequences that encode known proteins (see Table 5.6); 5 proteins were each encoded by the nucleotide sequence of 2 positive colonies (see Table 5.7); and there were 3 proteins that were encoded by the nucleotide sequence of three or more colonies (see Table 5.8). Due to the nature of a PhD, it was impossible to further investigate all of these proteins in detail; therefore I decided to only follow up those proteins identified 3 or more times.

The 3 proteins that were isolated from the library screens 3 or more times are called Homer3, Dncl1 (Dynein cytoplasmic light chain 1) and Trim11 (Tripartite motif protein family member 11). More information about the different clones that have been isolated for each of these proteins will be discussed in more detail in section 5.5 as well as information about the known structure and function of these proteins in section 5.6.

| Isolated from the C terminal domain library screen | |
|---|---|
| Gstm1 | Glutathione S transferase mu 1 |
| **Isolated from the PST domain library screen** | |
| Arl6ip5 | ADP-ribocylation like factor 6 interacting protein 5. |
| Snn | Aligned to the Stannin gene. |
| Nsg2 | Neuron specific gene family member 2. |
| Bcap37 | B cell receptor associated protein 37. |
| Arnt | Aryl hydrocarbon receptor nuclear translocator. |
| Tm9sf2 | Transmembrane 9 superfamily member 2. |
| Atp6f | Vacuolar proton-translocating ATPase 21kDa subunit. |
| Trb | Aligned to traube mRNA |
| ZfRbp | Zinc finger RNA binding protein |
| Vim | Vimentin is an intermediate filament cytoskeletal element expressed in brain, and also in retina and lens. Vim null mutations in mice do not affect mouse development. |
| Fbf1 | Structural protein with sequence similarity to plectin. May interact with CD95/Apo-1/Fas. |
| Pdcd6 | Programmed cell death 6 protein. |
| Cyt b | Cytochrome b from mitochondrial genome. |
| Fth | Ferritin heavy chain is a transport cargo protein that is cytoplasmic and expressed in brain. |
| mTRIP-Br1 | TRIP-Br are a novel family of PHD zinc finger and bromodomain interacting proteins. |
| Tap1 | Similar to transient receptor protein 4 (Trp4) associated protein. |
| Dncic1 | Cytoplasmic dynein intermediate chain 1B. |
| T-plastin | Similar to plastin 3 precursor, T plastin. |
| Mcmd7 | Mini chromosome maintenance deficient 7. |

**Table 5.6 Interactors isolated once that correspond to known proteins.**

The table above shows the proteins that were isolated once from the C terminal domain screen and the PST domain screen.

| Interactor | Description |
|---|---|
| Rab1 | Member of ras superfamily and possibly involved in the regulation of membrane vesicle targeting and fusion during organelle assembly and transport (Zahraoui *et al.*, 1989). |
| Cox1 | Cytochrome C oxidase subunit in mitochondrial genome. |
| Similar to KIAA1816 | Similar to the human predicted brain protein (Nagase *et al.*, 2001). |
| Trim7 and similar to GNIP2 | In humans TRIM7 is derived from GNIP which is a novel protein that binds and activates glycogenin and may represent a novel participant in the initiation of glycogen synthesis (Skurat *et al.*, 2002). |
| Ppia | Peptidylprolylisomerase A or cyclophilin A to be involved in protein folding and/or intracellular protein transport (OMIM 123840). |

**Table 5.7 Interactors isolated twice in the PST domain library screen.**

The following table outlines the interactors that were isolated twice in the PST domain library screen. The table gives the names and a very brief description of the proteins aligned to the sequence of positive colonies.

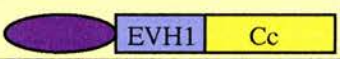
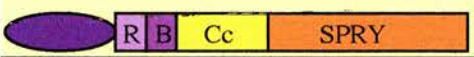
| Interactor | No. of clones | Isolated in library |
|---|---|---|
| HOMER3   | 7<br>6 | PST domain  <br>C terminal domain  |
| DNCL1  | 2<br>2 | PST domain  <br>C terminal domain  |
| TRIM11   | 6<br>0 | PST domain  <br>C terminal domain  |

**Table 5.8 Interactors isolated more than twice in the yeast 2-hybrid screens.**

The following table specifies the 3 interactors that were isolated more than twice in the yeast 2-hybrid screens. The table illustrates each of the isolated interactors including known domains (Cc, coiled coil domain, R, RING finger domain, B, B-box domain, EVH1 and SPRY see 5.6) coupled to the GAL4 activation domain (purple oval). The table also outlines the number of times each interactor was isolated from either the PST domain (ASNT) or the C terminal domain (TTST) yeast 2-hybrids screens (These constructs are also represented by schematic diagrams, see Table 4.1 for details).

## 5.5 Isolated Clone Information.

For each of the isolated clones that were identified for Homer3, Dncl1 and Trim11 the first thing that I looked at was whether these proteins would be expressed in the yeast 2-hybrid system. That is to say whether the coding regions for these proteins are in frame with the GAL4 DNA activation domain. It was confirmed that all of the 23 clones that encode these 3 proteins were in frame with this domain.

### 5.5.1 Homer3.

6 of the Homer3 clones were isolated from the TTS screens: 4 from the first screen as 'probably weak interactors' and 2 from the third screen as 'possible interactors or mixed populations'. 7 of the Homer3 clones were isolated from the ASN screen: 6 as 'interactors' (2 of which were from the high throughput X-gal assay) and one as a 'probable interactor'. When sequenced all the clones aligned to the sequence for the mouse, similar to Homer, neuronal immediate early gene, 3 (BC005773, IMAGE: 3602414). When the clones were compared 12 clones were exactly the same starting at nucleotide 334 and one clone (T24) contained an extra 3 nucleotides and started at nucleotide 331 in the published sequence of *Homer3*.

In the published sequence of *Homer3*, the normal starting methionine is at nucleotide position 120 and so the isolated clones do not code for the first 72 amino acids of Homer3 (T24 missing only the first 71 amino acids). Sequencing of the entire clones enabled me to discover that there was only a single nucleotide change in the coding region of the isolated clones from the published sequence and that this difference does not equate to an amino acid change. The stop codon of *Homer3* is at position 1188 in the published sequence and all of the sequenced clones retain this stop codon in the correct reading frame. All the clones extend up to the poly(A) stretch in the published sequence of *Homer3*, but the T24 clone has a slightly different end as the sequence AAATG is inserted into the beginning of this poly(A) stretch at nucleotide position 2373.

The isolated clones were shown to be missing the first 71/72 amino acids of the Homer3 protein, which corresponds to the N terminal half of the EVH1 domain (see 5.6.1) of the full length Homer3 protein. I was initially concerned that the clones devoid of these amino acids would have a completely different structure. By using the secondary structure prediction program JPRED (see 2.1.3) I showed that the missing N terminal 71/72 amino acids do not drastically alter the secondary structure of the rest of the protein, and that the missing region is composed of

β sheets bordered by 2 alpha helixes. Nevertheless, I decided that it would be beneficial to create a full length *Homer3* cDNA (see 6.2) to investigate any difference to the truncated Homer3 protein in pairwise interactions with PAX6 (see Chapter 6).

### 5.5.2 Dncl1 (Dynein cytoplasmic light chain 1).

2 of the Dncl1 clones were isolated from the second TTS screen high throughput X-gal assay as 'possibly weak interactors'. The remaining 2 Dncl1 clones were isolated from the ASN screen as 'interactors'. The sequence alignment of these 4 clones was to the nucleotide sequence for the mouse dynein, cytoplasmic light polypeptide clone (BC008106, IMAGE: 3158148). In this sequence of *Dncl1* the normal starting methionine is at nucleotide position 99 and the stop codon at position 366. The coding region of the 4 isolated clones is exactly the same as the published sequence as is the 3' UTR which extends for a further 273 nucleotides before the poly(A) stretch.

However, there is slightly more variability between the clones in the 5' UTR. One clone (XT31) is exactly the same as the published nucleotide sequence with an extra 98 nucleotides before the starting methionine; one clone (A13) only has an extra 71 nucleotides before the start codon; one clone (XT84) has an extra 56 nucleotides before the start codon; the final clone (A62) has the same 56 nucleotides before the start codon as XT84, however, it has an additional 48 nucleotides inserted 8bp before the start site of the protein. The inserted nucleotides of clone A62 begin with the sequence GTGAGA that resembles the consensus splice donor sequence (GTAAGT) (Krawczak *et al.*, 1992) and so this clone could possibly represent a splice variant of *Dncl1*.

### 5.5.3 Trim11 (Tripartite motif protein family member 11).

All 6 of the Trim11 clones were isolated from the ASN screen as 'interactors' (2 of which were from the high throughput X-gal assay). When sequenced all the 6 isolated clones aligned to the nucleotide sequence of the mouse tripartite motif protein 11 (Trim11, AF220124). There are no differences between the coding region of the isolated clones and the published nucleotide sequence that has a starting methionine at position 125 and a stop codon at position 1526. All the 6 clones have the same 3' UTR as the published sequence which extends for a further 723 nucleotides after the stop codon before the poly(A) stretch. However, there is more variation in the length of the 5' UTR; 2 clones (A3, A5) have 92 nucleotides; 3 clones (A23, A47, XA22)

have 125 nucleotides, which is the same as in the published sequence; and 1 clone (XA3) has 143 nucleotides at the N terminus, which is 18bp more than in the published sequence.

## 5.6 Interacting protein information.

In this section I have reviewed the current information about each of these potential novel and interesting interactors. I have included information about the structure and function of each of the proteins and at the end of each section have included some hypotheses of what role the interactions with PAX6 may play in development. In the sections that follow (5.6.1 to 5.6.3) for the sake of simplicity I will use generic lower case gene and protein names, even when I am referring to the human gene or protein.

### 5.6.1 Homer3.

In *Drosophila* there is only a single *Homer* gene that encodes a single Homer protein, however, in mammals there are three *Homer* genes, *Homer1*, *Homer2* and *Homer3*, which encode for several Homer protein isoforms (Xiao *et al.*, 1998). The *Homer3* gene is located on chromosome 19q13.11 in humans and 8B3.3 (69.3cM) in mice. The human and mouse genes are thought to be composed of 10 exons that encode a single protein isoform of 365 amino acid and 357 amino acid proteins respectively (www.ensembl.org). There is no co-localisation between any interesting disease phenotypes and the chromosomal location of *Homer3* when investigating the OMIM database and the mouse genome informatics (MGI) database (www.informatics.jax.org/).

The initial Homer protein that was characterised in humans was the Homer1a isoform that is highly up regulated by synaptic activation and is only composed of the N terminal EVH1 domain (Brakeman *et al.*, 1997; Kato *et al.*, 1998). However, the Homer3 protein, as well as other isoforms such as Homer1b/c, and Homer2a/b are composed of 2 domains: the EVH1 domain encompassing the N terminal 186 amino acids in mice; and a coiled coil domain produced from 2 leucine zipper motifs (Sun *et al.*, 1998) within the C terminal 171 amino acids in mice (see Table 5.8). The EVH1 domain is highly conserved between the members of the Homer protein family, for example in the N terminal 120 amino acids of the EVH1 domain, Homer1a and Homer3 are 86% homologous (Xiao *et al.*, 1998). However, this homology is vastly reduced between the C terminal sequences of Homer3 and the long isoforms Homer1b/c with only 22% homology, although even with this reduced homology the coiled coil secondary

structure is retained between all of the Homer proteins (Xiao *et al.*, 1998). Specifically in mice and humans the Homer3 protein sequence is 88.6% identical.

The N terminal region of Homer proteins was initially thought to be homologous with a PDZ domain, which is involved in protein-protein interactions and PDZ domain containing proteins are known to be enriched in the post synaptic density (PSD) (Brakeman *et al.*, 1997). The PSD is a region subjacent to the postsynaptic membrane of excitatory synapses, predominantly found on dendritic spines (specialised structures that protrude from the dendritic shaft), which often contain smooth endoplasmic reticulum and translational machinery (Sala *et al.*, 2001). However, it has been suggested that there is a closer homology to the EVH1 domain (Kato *et al.*, 1997). The EVH1 domain was initially characterised due to its similarity to the Ena (Enabled), VASP (vasodilator stimulated phosphoprotein) or WASP (Wiskott-Aldrich Syndrome homology domain) protein domains (Kato *et al.*, 1997) that are common to cytoskeleton interacting proteins (De Bartolomeis and Iasevoli, 2003). EVH1 domains are found in proteins that range in function from cell signalling, nuclear transport and cytoskeletal events (Xiao *et al.*, 1998) and this domain binds to a consensus polyproline motif, FPxxP (Ball *et al.*, 2002). However, the Homer EVH1 domain (from this point on merely referred to as the EVH1 domain) has been proposed to be a novel functional subset of classical EVH1 domains, as crystallographic studies revealed that this domain has a unique structure which allows binding to a different but specific polyproline motif xPPxxF (Beneken *et al.*, 2000).

The specific polyproline motif (xPPxxF) that is bound by the EVH1 domain of Homer3 is not found in the PST domain sequence of Pax6 and although a PPMPSF sequence is in Pax6, when compared to the specificity of the motif in other binding partners of Homer3 (Ball *et al.*, 2002) I would suspect that this does not represent a true Homer3 binding site. Furthermore, 3 surface-exposed aromatic side chains are vital for the recognition of proline rich sequences in ligands by the EVH1 domain (Ball *et al.*, 2002) and 2 of these are within the first 70 amino acids of Homer3 which is missing in the Homer3 clones isolated from the yeast 2-hybrid screen. This seems to suggest that Pax6 interacts with the coiled coil domain of Homer3, one of the few interactions that have been reported associating with this region of the Homer3 protein. Obviously more experimentation must be performed to elucidate the site of interaction between Homer3 and Pax6 more specifically.

The coiled coil domain is produced by multimerisation of the 2 leucine zipper motifs within the C terminal region of Homer3 (Sun *et al.*, 1998). These leucine zipper motifs are known to

promote self association; however it has been shown that this region can also facilitate multimerisation between the coiled coil domains of different members of the Homer protein family producing heteromeric clusters (Shiraishi *et al.*, 2003).

The Homer family can be divided into 2 groups: the inducible form, Homer1a that can be induced by synaptic stimuli; and the constitutively expressed forms, Homer1b/c, Homer2 and Homer3. The Homer3 protein has been shown to be expressed in the brain but more specifically in the hippocampus and the cerebellum (Xiao *et al.*, 1998) as well being found in the lung and thymus (Xiao *et al.*, 1998). Homer3 is located in the postsynaptic density, as well as being distributed along the soma and into apical dendrites, but is not present in the cell nucleus (Xiao *et al.*, 1998). Interestingly it has been shown that Homer3 is expressed in discrete puncta along the length of dendritic shafts in hippocampal neurons in mammals (Shiraishi *et al.*, 2003) and *Drosophila* (Diagana *et al.*, 2002) and this expression pattern of Homer proteins overlap with synaptogenesis in development. Therefore, it has been suggested that Homer proteins are involved in the formation of synapses (Kato *et al.*, 1998; Diagana *et al.*, 2002) as well as possibly being involved in long term potentiation in excitatory synapses with the possibility of functioning in memory formation (Kato *et al.*, 1997).

The EVH1 domains of Homer proteins, and more specifically Homer3, have been shown to interact with the polyproline motifs of: type-I metabotropic glutamate receptors (mGluRs), mGluR1α and mGluR5 (Brakeman *et al.*, 1997); inositol 1,4,5-trisphosphate receptors (IP3Rs) (Tu *et al.*, 1998); ryanodine receptors (RyRs) (Kato *et al.*, 1998); and the scaffolding protein Shank (Tu *et al.*, 1999) (see Figure 5.4). However, it must be noted that until recently no proteins have been reported that interact with the coiled coil domain of Homer proteins.

It has just been reported that the N terminus of the human S8 ATPase, which has a role in proteasomal processing of proteins for degradation, interacts with the coiled coil domain of Homer3 (Rezvani *et al.*, 2003). Homer3 binds to the C terminal region of mGluR1α, and simultaneous binding of the S8 ATPase has been shown to be involved in Homer3 mediated degradation of these receptors (Rezvani *et al.*, 2003).

The EVH1 domain of Homer3 interacts with the C terminal tail of mGluR1α and mGluR5 (Tu *et al.*, 1998), which have been linked to determining synaptic plasticity, neuronal development and pathological conditions such as epilepsy or Parkinson's disease due to their involvement in multiple transduction pathways (De Bartolomeis and Iasevoli, 2003). However, the EVH1

domain of Homer3 interacts with the cytosolic N terminal regions of the IP3Rs and RyRs, which are known to be involved in calcium release due to their link with the endoplasmic reticulum (ER) (Tu *et al.*, 1998). Therefore, due to the multimerisation via the coiled coil domain the Homer3 protein is able to link these 2 types of receptor with the mGluRs (see Figure 5.4), which has been functionally demonstrated with Homer3 co-precipitating with IP3R and mGluR1α from cerebellar lysates (Tu *et al.*, 1998).

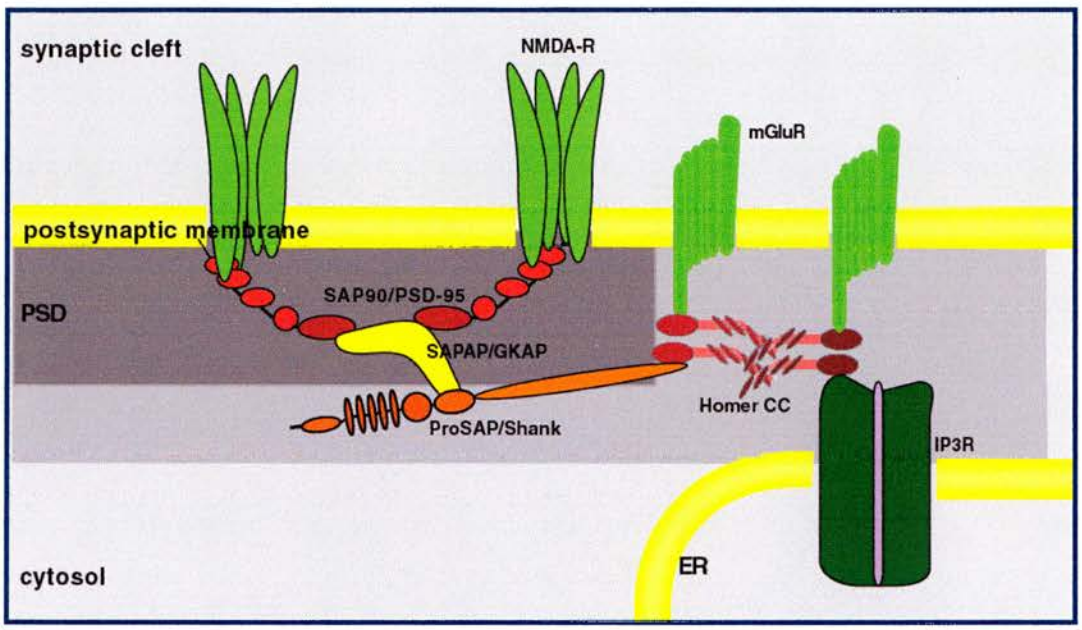



**Figure 5.4 Role of Homer3 as a scaffolding protein in the PSD density.**

This illustrates how the Homer proteins (Homer CC), including Homer3 multimerise through their coiled coil regions and through their EVH1 domain interact with: type-I metabotropic glutamate receptors (mGluRs), mGluR1α and mGluR5; inositol 1,4,5-trisphosphate receptors (IP3Rs); and the scaffolding protein Shank. The interaction between Homer CC and Shank provides a link between the mGluRs and the N-methyl-D-aspartate receptors (NMDA-R) in the post-synaptic density (PSD) *(Thomas, 2002)*.

This scaffolding function of Homer3 as well as the other long Homer isoforms (i.e. those which contain both EVH1 and coiled coil domains) is antagonised by Homer1a in a dominant negative way (Xiao *et al.*, 1998) (see Figure 5.5). As already mentioned the expression of Homer1a is rapidly induced by synaptic activity (Brakeman *et al.*, 1997; Kato *et al.*, 1997; Ango *et al.*, 2001) and this disrupts the complex between Homer3 and both mGluRs and IP3Rs which results in an alteration of the mGluR induced calcium release (Tu *et al.*, 1998). The disruption of the complex between Homer3 and mGluR1α by Homer1a has specifically been shown in cerebellar homogenates (Xiao *et al.*, 1998).

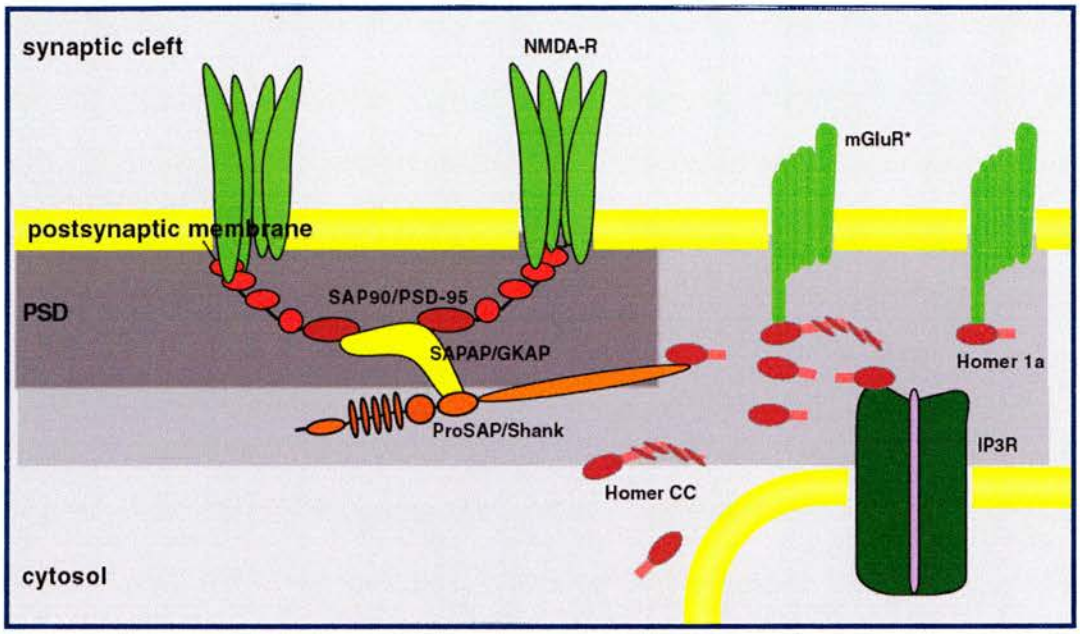



**Figure 5.5 Scaffolding function of Homer3 antagonised by Homer1a.**

This illustrates how the scaffolding function of Homer proteins (Homer CC), including Homer3 is antagonised by Homer1a. Homer1a disrupts the complex between Homer CC and: type-I metabotropic glutamate receptors (mGluRs), mGluR1α and mGluR5; inositol 1,4,5-trisphosphate receptors (IP3Rs); and the scaffolding protein Shank. Uncoupled mGluR* may have agonist-independent activity *(Thomas, 2002)*.

It has been shown that Homer1a mRNA expression is induced in the cerebral cortex of dark-reared rats 30mins after their first visual experience in the 3-5 postnatal weeks (Brakeman *et al.*, 1997). Also monocular deprivation by stopping the retinal activity of one eye means that the mRNA levels of Homer1a are rapidly reduced in the contralateral visual cortex (Brakeman *et al.*, 1997). Therefore, this provides an example of how physiological neuronal activity dramatically regulates Homer1a, which reinforces the suggestion that the Homer proteins are involved in stimulating synapse formation (Brakeman *et al.*, 1997).

It was also found that the mGluR1α was able to produce agonist-independent activity when Homer3 levels were decreased in cerebellar granule cells (Ango *et al.*, 2001). This was investigated further and it was found that the cessation of the constitutive activity of mGluR1α required not only the binding of Homer3 to the receptor but also the coupling of the receptor via Homer3 to other synaptic proteins such as Shank (Ango *et al.*, 2001). Therefore, this provides a hypothesis in which the competitive interaction of Homer1a with the receptors promotes the constitutive activity of the mGluR1α due to its uncoupling from the synaptic complex without

receptor activation (Ango *et al.*, 2001) (mGluR*, see Figure 5.5). This mechanism might be utilised in situations such as synaptic plasticity or in seizures (Ango *et al.*, 2001; Thomas, 2002) where a delayed or long lasting signal is produced from an initially strong stimulus.

The interaction between the EVH1 domain of Homer proteins and Shank proteins (see Figure 5.4) has been shown to occur in mammals (Tu *et al.*, 1998) and in *Drosophila* (Diagana *et al.*, 2002) showing that this interaction is evolutionarily conserved, although a specific interaction with Homer3 can only be alluded to due to EVH1 domain homology. It has been suggested that an important function of Shank is to stimulate head expansion in dendritic spines and that this role is reliant on the interaction between Homer and Shank (Sala *et al.*, 2001; Ehlers, 2002). As a consequence of this spine enlargement the occurrence of other synaptic proteins was increased, most notably IP3Rs and F-actin, which have both been shown to interact with Homer (Xiao *et al.*, 1998; Shiraishi *et al.*, 1999). It was suggested that Homer mediates their enrichment in the PSD due to its interaction with Shank (Sala *et al.*, 2001). Therefore, the Homer/Shank complex seems to indicate that these proteins are perhaps involved in developmental and activity dependant structural plasticity (Thomas, 2002), which has been suggested to underlie experience dependant changes in the brain such as learning and memory (Ehlers, 2002).

As I have already mentioned Homer proteins function as scaffolding proteins within the PSD of dendritic spines, interlinking proteins involved in synaptic transmission. Further evidence for this has been shown as Homer proteins were colocalised with the NMDA receptor complex during dendritic and synaptic differentiation in excitatory hippocampal neurons (Shiraishi *et al.*, 2003). This colocalisation has been suggested to be due to the interaction between Homer and Shank, which seems to be part of a larger complex involving the binding of Shank and GKAP and PSD-95 as well (Grant and O'Dell, 2001). It is thought that this complex may function to make a functional bridge between mGluRs and NMDA receptors (see Figure 5.4) in neuronal development and synaptic plasticity (Brakeman *et al.*, 1997; Tu *et al.*, 1999).

*Drosophila* mutants of the single Homer protein have been shown to be predominantly normal with no abnormalities in axon pathfinding during development, no synaptic transmission problems or general nervous system development problems and these mutant flies are viable, fertile and can respond to visual stimuli (Diagana *et al.*, 2002). However, it has been reported that over expression of either Homer1a or Homer1b/c in *Xenopus* optic tectal neurons resulted in aberrant axon pathfinding due to a dysfunction of the signal transduction from cell surface receptors to intracellular proteins (Foa *et al.*, 2001). However, a gain of function Homer

mutation in *Drosophila* did not show any axon guidance problems, which is consistent with the finding that in *Drosophila* little or no Homer protein is detected in the axons of neurons, as the expression is concentrated in the dendrites of neuronal cells (Diagana *et al.*, 2002). *Drosophila* Homer mutants showed defects in behavioural plasticity. They were unable to learn/memorise mating conditioning. Also they showed defects in the control of locomotor activity, characterised by higher levels of spontaneous locomotor activity (Diagana *et al.*, 2002). Although these Drosophila mutants did not confirm a role for Homer in synaptogenesis it was suggested that Homer predominantly functions to alter the activity of neuronal circuits by regulating synaptic functioning due to its control of mGluRs signalling (Diagana *et al.*, 2002), backed up by the fact that mice with a loss of function mutation of type-I mGluRs also develop learning problems as well as movement problems (Aiba *et al.*, 1994).

With the importance of the type-I mGluRs clearly evident in correct synaptic functioning it is interesting to note that Homer is central to their function in protecting neurons from apoptotic death (Rong *et al.*, 2003). I have already mentioned how coupling of S8 ATPase to mGluRs via Homer3 targets these receptors for degradation, however, it has recently been shown that the activity of the mGluRs promoted the recruitment of a phosphoinositide 3-kinase enhancer (PIKE-L) to the receptor due to its interaction with Homer. This has the effect of activating phosphoinositide 3-kinase (PI3 kinase), which had the effect of preventing neuronal apoptosis (Rong *et al.*, 2003).

I have already explained how the interaction between Shank and Homer is central to the formation of a complex that makes a functional bridge between mGluRs and NMDA receptors in neuronal development and synaptic plasticity (Brakeman *et al.*, 1997; Tu *et al.*, 1999). Furthermore, the activation of NMDA receptors stimulates mitogen-activated protein kinase (MAPK), which results in long term potentiation (LTP). LTP and long term depression (LTD) are both changes in the strength of synapses that are caused by changes in their activity and both LTP and LTD have been implicated in memory and learning (Grant and O'Dell, 2001).

MAPK has been shown to be physically tethered to the NMDAR/PSD-95 complex and thought to act on some immediate synaptic substrates, which allows LTP induction (Grant and O'Dell, 2001). MAPK has also been implicated in phosphorylating Pax6 and it has been suggested that MAPK is involved in post-translational regulation of Pax6 after activation of MAPK signalling cascade (Mikkola *et al.*, 1999). Therefore, my discovery of an interaction of Pax6 and Homer3 is intriguing as it is possible that this interaction establishes a physical proximity between Pax6 and

MAPK in the PSD of neurons. Therefore, the activation of MAPK by NMDAR stimulation could lead to a rapid phosphorylation of Pax6, which cause an increase in the transcriptional activation by Pax6 (Mikkola *et al.*, 1999). Therefore, it is possible that this increased transcriptional activation may function in the LTP, learning, memory or even spine morphogenesis in the developing brain to develop new synaptic connections.

A huge yeast 2-hybrid screen was performed by the CuraGen Corporation, which investigated potential proteins interactions of the entire proteome of *Drosophila melanogaster* (Giot *et al.*, 2003). The results of the yeast 2-hybrid screen were then designated as being either 'high confidence' interactions or 'other interactions' and can be searched at the *Drosophila* interaction database (http://portal.curagen.com/cgi-bin/interaction/flyHome.pl?modeIn=List). This revealed that in *Drosophila* the Pax6-like protein eyegone (eyg) interacts with the *Drosophila* Homer protein. The eyg protein appears to corresponds to Pax6 missing the NTS subdomain of the PD, which is the equivalent of the +5a Pax6 isoform in vertebrates (Epstein *et al.*, 1994).

### 5.6.2 Dncl1 (Dynein cytoplasmic light chain 1).

The *Dncl1* gene is located on chromosome 12q24.31 in humans and 5F (112.5cM) in mice and the human gene is composed of 3 exons encoding a protein of 89 amino acids (www.ensembl.org). There is 89% identity between the mouse and human nucleotide sequences but the protein sequence is identical between mouse and humans as well as being 90-100% conserved with homologous proteins from mammals, insects, nematodes and bacteria and 50-60% conserved in yeast and plants (Jacob *et al.*, 2000). It is interesting to note that by looking at the OMIM gene database it can be seen that the 12q22-q24.1 region co-localises with growth and mental retardation, and the 12q24.1q31-q42 region co-localises with Schizophrenia. There is no co-localisation with any interesting phenotypes when looking at the mouse genome informatics (MGI) database (www.informatics.jax.org/).

Whilst reading the literature on the dynein cytoplasmic light chain 1, I realised that there were many synonyms for Dncl1 in the literature that included LC8, PIN, DLC1 and DLC8. However, in an effort to simplify things I have chosen to merely use the synonym Dncl1/ DNCL1.

Dncl1 is a component of the brain cytoplasmic dynein complex (King *et al.*, 1996) (see Figure 5.6), which is a molecular motor involved in minus end directed movement along microtubules i.e. in neurons this movement is from the synapse to the neuronal cell body (Jacob *et al.*, 2000)

(see Figure 5.7). Not only is Dncl1 thought to be a component of the dynein motor complex, but there is also evidence that it is a component of an actin based molecular motor (Espindola *et al.*, 2000). A protein homologue of Dncl1 has been reported in chicken associated with the tail domain of chicken brain myosin-Va (see Figure 5.6) which is a ubiquitously expressed actin based molecular motor (Espindola *et al.*, 2000). Therefore, Dncl1 is thought to be a subunit of both a microtubule transport network involved in moving protein cargos large distances and an actin transport network which is responsible for the local distribution of proteins within the cell (Jacob *et al.*, 2000).

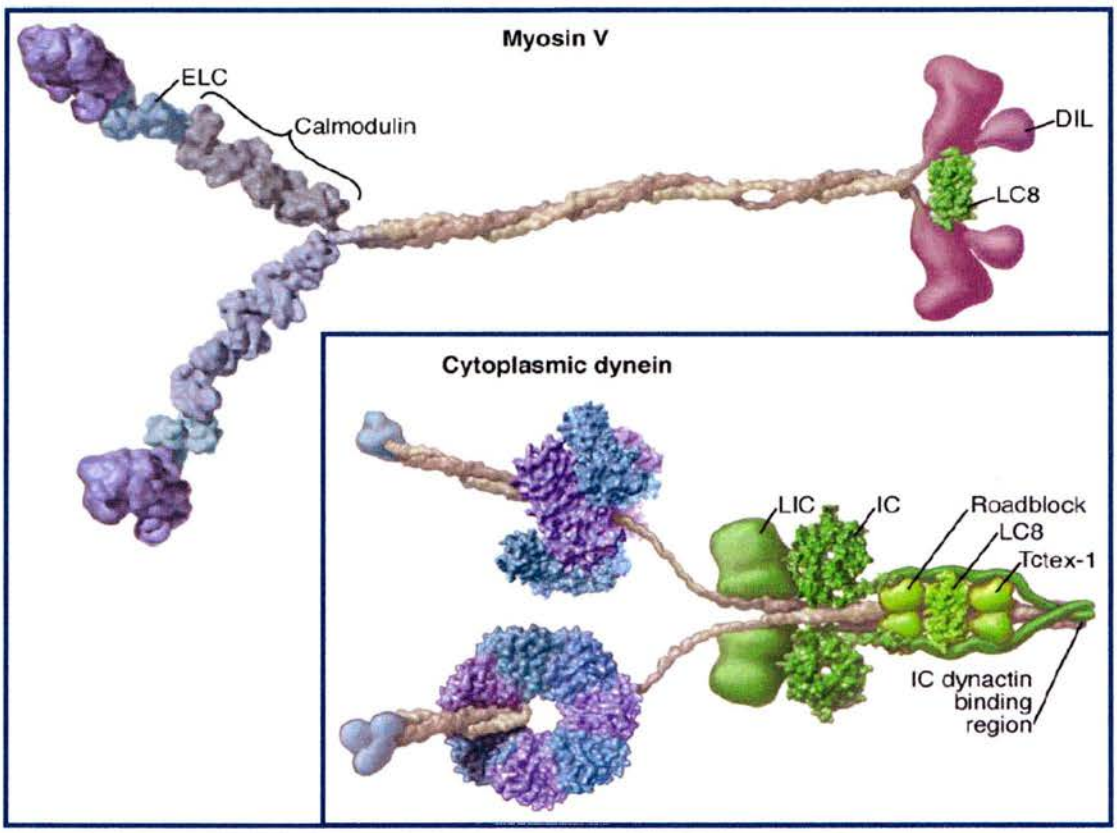



**Figure 5.6 Dncl1 incorporation in motor protein complexes.**

This illustrates how Dncl1 (LC8) is incorporated into the actin based (Myosin V) and microtubule based (Cytoplasmic dynein) molecular motor proteins (Vale, 2003). Nomenclature: DIL represents the dilute domain; ELC is the essential light chain; LIC is the light intermediate chain; and IC is the intermediate chain.

Dncl1 was originally described as being homogeneously expressed in the cytoplasm of cells (Dick *et al.*, 1996; Jaffrey and Snyder, 1996; King *et al.*, 1996), however more recently it has been reported that Dncl1 is present in both the nucleus and the cytoplasm (Crepieux *et al.*, 1997;

Herzig *et al.*, 2000). This difference in the subcellular localisation of Dncl1 has been ascribed to differences in cell type and cell cycle phase (Kaiser *et al.*, 2003).

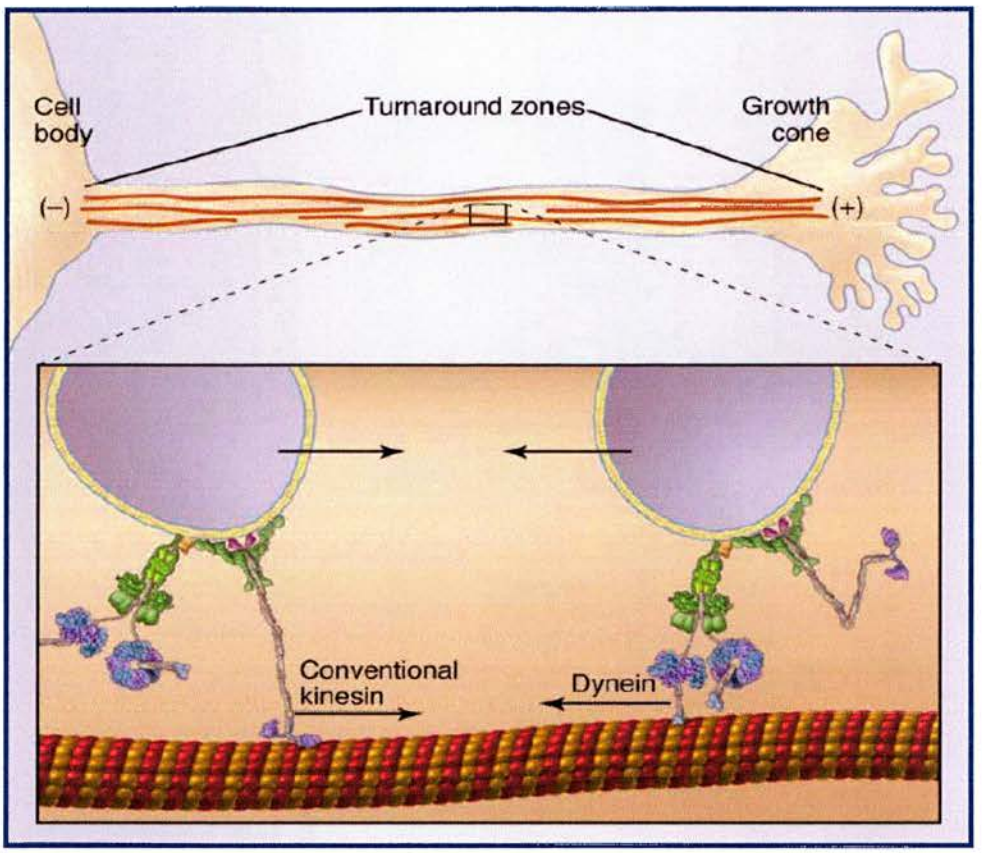



**Figure 5.7 Minus end directed movement of dynein along microtubules.**

This illustrates how dynein binds to cargo proteins via Dncl1 and allows retrograde transport from the growth cone or post-synaptic density to the nucleus (Vale, 2003).

Nuclear and cytoplasmic extracts of two mammalian cell lines showed that Dncl1 was expressed in both the nucleus and cytoplasm, which led to the hypothesis that Dncl1 may be involved in intracellular trafficking of protein complexes (Crepieux *et al.*, 1997). Furthermore, it has been suggested that perhaps the binding of Dncl1 to proteins facilitates their nuclear export and subsequent cellular distribution due to coupling to the microtubule organising centre (MTOC), which is a perinuclear region from which the microtubule network radiates (Crepieux *et al.*, 1997). This hypothesis is based on the observation that Dncl1 has been shown to colocalise with both IκBα (Crepieux *et al.*, 1997) and the African swine fever virus (ASFV) p54 protein (Alonso *et al.*, 2001) at the MTOC.

Dncl1 has also been reported to occupy discrete regions of the nucleus colocalising with the transcriptional repressor protein (TRPS1) (Kaiser *et al.*, 2003) and the transcriptional activator nuclear reparatory factor 1 (NRF1) (Herzig *et al.*, 2000). This has led to the suggestion that Dncl1 may be involved in the movement of proteins such as transcription factors from interchromatin granules (storage or assembly structures for proteins involved in gene expression) in the nucleus to sites of active gene expression (Herzig *et al.*, 2000).

Dncl1 has also been shown to bind to specific sequences in the 3'UTR of mRNA of the parathyroid hormone (PTH). It has been proposed that this interaction facilitates a specific cytoplasmic localisation of the mRNA, due to its coupling to the microtubule associated dynein motor complex, to produce a spatially restricted pattern of protein expression in the cell (Epstein *et al.*, 2000).

In situ hybridisation using *Dncl1* cDNA showed that it was strongly expressed in the central nervous system of Drosophila (Herzig *et al.*, 2000) and high concentrations of *Dncl1* mRNA have been shown to be present in brain, heart, liver and kidney by Northern blot analysis in rat (Wilson *et al.*, 2001). Furthermore, it has been shown by immunofluorescence that Dncl1 is a synaptic protein that is predominantly concentrated at the postsynaptic density (PSD) but is also expressed in a punctate pattern along dendritic shafts (Naisbitt *et al.*, 2000).

Dncl1 has also been shown under the synonym of pin (protein inhibitor of neuronal nitric oxide synthase) to be an inhibitor of nNOS (neuronal nitric oxide synthase) (Jaffrey and Snyder, 1996). It was reported that binding of the Dncl1 protein to nNOS destabilised the dimeric structure, which is required for the enzyme's activity (Jaffrey and Snyder, 1996). However, more recently this role of Dncl1 has been challenged and it has been proposed that the main function of Dncl1 is in dynein and myosin motor complexes (Hemmens *et al.*, 1998). This is further substantiated as it has been demonstrated that Dncl1 definitely binds to nNOS; however, it neither dissociates the dimeric structure of the enzyme nor inhibits its function. Therefore, it is proposed that its interaction merely facilitates the transport of nNOS along the microtubule network (Rodriguez-Crespo *et al.*, 1998).

Not only has Dncl1 been shown to bind to nNOS (Rodriguez-Crespo *et al.*, 1998) but it has also been shown to bind to the guanylate kinase domain-associated protein (GKAP) (Naisbitt *et al.*, 2000). In both cases Dncl1 has been proposed to function in coupling these proteins to cytoplasmic dynein and so facilitating their axonal transport. Interestingly, the post-synaptic

density (PSD), a dynamic structure where proteins are transient and stoichiometrically variable (Sheng and Kim, 2002), has been shown to include both nNOS and GKAP coupled to NMDA (*N*-methyl-D-aspartate) receptors through the PSD-95 family of scaffold proteins (see Figure 5.4) (Sheng and Kim, 2002). Therefore, it is possible that both nNOS and GKAP are transported to the PSD and due to their interaction with Dncl1 are specifically localised within this region via the myosin molecule, as it has also been shown that PSD-95, GKAP, Dncl1 and myosin-V co-immunoprecipitate as a complex in rat brain (Naisbitt *et al.*, 2000). This is fascinating as it is possible that a similar function of Dncl1 could be attributed to its interaction with PAX6 and that Dncl1 is responsible for specifically localising PAX6 within the PSD in developing neurons.

The possible roles of the Dncl1 protein that I have mentioned so far have been in intracellular transport and more specifically in axonal transport; however, other functions of the Dncl1 protein have also been described including: regulation of the pro-apoptotic activity of Bim, a Bcl2 family member (Puthalakath *et al.*, 1999); and suppression of the transcriptional repression function of TRPS1 (Kaiser *et al.*, 2003). In the case of TRPS1 the suppression of transcriptional repression is thought to be due to a possible conformational change due to the interaction of TRPS1 and Dncl1 (Kaiser *et al.*, 2003). However, the Bim protein is regulated as it is attached to the microtubule associated dynein motor complex via Dncl1. When activated by certain apoptotic stimuli the Bim/Dncl1 complex dissociates from the dynein molecule and moves to the Bcl2 protein where it interacts and stops the antiapoptotic effect of Bcl2 protein (Puthalakath *et al.*, 1999).

Loss of function mutations in *Drosophila* show embryonic lethality via the apoptotic pathway (Dick *et al.*, 1996); however, this is not surprising given that Dncl1 functions to sequester the proapoptotic protein Bim onto the microtubule network (Puthalakath *et al.*, 1999). Although lethal this mutation is characterised by severe neuronal defects suggesting that Dncl1 may also be involved in axon formation (Jacob *et al.*, 2000).

Whilst reading the literature on Dncl1 I discovered that potential Dncl1 binding sites in the interacting proteins had been ascertained in many cases. Therefore, I thought it would be interesting to use matcher in EMBOSS (see 2.1.5) to identify any regions of similarity between these Dncl1 binding sites and the sequence of the PST domain of PAX6. As I have illustrated in Figure 5.8 there are potentially two Dncl1 binding domains within the PST domain of PAX6, although this hypothesis is based on very weak regions of homology between previously isolated Dncl1 binding sites.

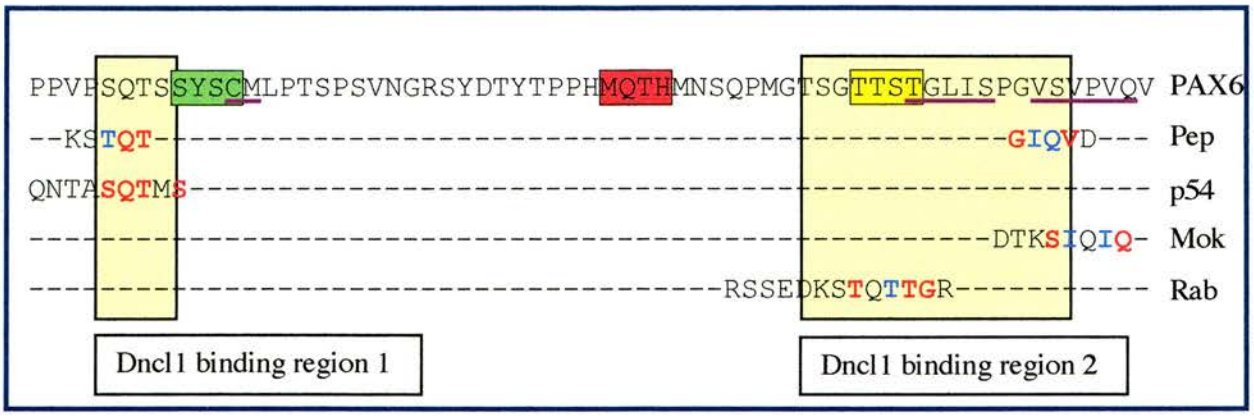



**Figure 5.8 Alignment of Dncl1 binding domains to PAX6 sequence.**

This shows the PAX6 amino acid sequence (345-407aa) as illustrated in Figure 3.1, as well as the alignment to the published Dncl1 binding domains that have previously been reported: as determined from pepscan technique (Pep) (Rodriguez-Crespo *et al.*, 2001); in the African swine fever virus (ASFV) protein p54 (Alonso *et al.*, 2001); in the Mokola virus (Mok) P protein (Jacob *et al.*, 2000); and in the Rabies virus (Rab) P protein (Poisson *et al.*, 2001). A **red** amino acid indicates an identical residue whereas a **blue** amino acid indicates a similar residue. Two possible Dncl1 binding regions in PAX6 have also been illustrated.

The two consensus sequences identified by Rodriguez-Crespo *et al.*, (2001) using a pepscan technique in Dncl1 interacting proteins are KSTQT and GIQVD, which correspond to Dncl1 binding regions 1 and 2 respectively in Figure 5.8. The interacting proteins investigated in this technique included some of the proteins I have already mentioned including the Bim protein (Puthalakath *et al.*, 1999), which has a KSTQT consensus sequence and nNOS and GKAP which both utilise the GIQVD sequence to bind to Dncl1 (Rodriguez-Crespo *et al.*, 2001). The sequence in the p54 protein of the African swine fever virus (ASFV) that has been shown to interact with Dncl1 shows similarity to Dncl1 binding region 1 (Alonso *et al.*, 2001). Human DNCL1 interacts with the phosphoproteins of 2 lyssaviruses: the rabies virus (Poisson *et al.*, 2001) and the Mokola virus (Jacob *et al.*, 2000). The sequences in both these virus P proteins that interact with Dncl1 show similarity to Dncl1 binding region 2.

Although the characterisation of these two potential Dncl1 binding regions in Pax6 was based on very weak regions of homology with other identified Dncl1 binding domains it is encouraging to see that the C terminal domain is potentially important in facilitating the interaction between Pax6 and Dncl1.

### 5.6.3 Discussion of the interaction of Homer3 and Dncl1 with Pax6.

Interestingly, it has been demonstrated that the localisation of the S8 ATPase in the dendrites of Purkinje cells is dependant on the presence of mGluR1α, otherwise the S8 ATPase is found in the nucleus of the neuron (Rezvani *et al.*, 2003). This suggests that the scaffolding function of Homer3 allows proteins that bind to the coiled coil domain of Homer3 to be specifically sequestered in the postsynaptic density. Therefore, the interaction of Homer3 and Pax6 could suggest that Pax6 is sequestered in the PSD until stimulus activation of Homer1a causes the dissociation of the Homer scaffolding proteins, including Homer3, so that Pax6 could translocate to the nucleus. Alternatively, stimulation of receptors in the postsynaptic membrane might also activate MAPK, as I have already mentioned, which could potentially activate Pax6 before the PSD complex of Homer3 and Pax6 dissociates.

It is even more intriguing to then include Dncl1 in this picture, so that the release of activated Pax6 from the PSD scaffold allows Pax6 to be translocated to the nucleus specifically by its interaction with Dncl1. This would mean that PAX6 could travel along the actin cytoskeleton first until it reached the microtubule network as Dncl1 has been shown to interact with both Myosin-Va and Dynein motor complexes. This provides a possible hypothesis for how the stimulation of synapses by extracellular cues that are detected by cell-surface receptors could lead to changes in gene expression within the cell nucleus of developing neurons, which could have implications for synapse function such as synaptic plasticity or long term potentiation. Obviously this is only hypothetical and the interactions need to be definitively confirmed.

### 5.6.4 Trim11 (Tripartite motif protein family member 11).

The *Trim11* gene is located on chromosome 1q42.13 in humans and 11B1.3 (59.6cM) in mice. The human and mouse genes are thought to be composed of 6 exons that encode 482 amino acid and 467 amino acid proteins respectively (www.ensembl.org), which show 90.5% similarity (86.6% identity). It is interesting to note that by looking at the OMIM gene database it can be seen that the 1q31-q42 region co-localises with Alzheimer's disease 4, whereas region 1q42.1 co-localises with schizophrenia and more specifically the *DISC1* and *DISC2* genes. There is no co-localisation with any interesting phenotypes when looking at the mouse genome informatics (MGI) database (www.informatics.jax.org/).

Trim11 is a member of the mouse tripartite motif protein family (also known as the RBCC family), and contains the characteristic structural motifs of this protein family: a RING finger, a

B-box zinc finger, and a coiled coil domain, as well as a B30.2 domain (also termed a SPRY domain) that is in many but not all Trim proteins (Reymond *et al.*, 2001).

The expression of Trim11 in mouse E12.5 embryos is ubiquitous, however, there is a stronger signal in the CNS, more specifically in the dorsal root ganglia, the telencephalon, the mesencephalon, and the external layers of the cerebellar primordium (Reymond *et al.*, 2001). Trim11 has a diffuse nuclear and cytoplasmic localisation and is not localised to specific cell compartments like other Trim proteins (Reymond *et al.*, 2001).

The RING finger is a 40 to 60 amino acid domain that is thought to be involved in protein-protein interactions (Saurin *et al.*, 1996). In acute promyelocytic leukaemia the promyelocytic leukaemia protein (PML) is fused to the retinoic acid receptor due to a chromosomal translocation (Topcu *et al.*, 1999). The PML protein is also a RBCC/TRIM protein and has been shown to form macromolecular complexes in the nucleus, called nuclear bodies (Saurin *et al.*, 1996) which are disrupted if the RBCC/TRIM domain (Borden, 2000) or more specifically, the RING domain (Saurin *et al.*, 1996) is mutated. This suggests that the RING domain is important in protein interactions and this is further supported by the observation that the RING domain of PML binds to a specific 65 amino acid sequence in the transcriptional repressor proline rich homeodomain (PRH) protein (Topcu *et al.*, 1999). The binding of the RING domain of PML increases the transcriptional repression of PRH, which suggests that this domain may be linked to a role in transcriptional repression. This is corroborated as the RING domain of the RBCC/TRIM protein KAP1 has been shown to bind to the KRAB domain of the transcriptional repressor KOX1 and regulate transcriptional repression through this interaction (Borden, 2000).

The RING domain is also thought to be involved in protein ubiquitination and function as an E3 ubiquitin ligase (Borden, 2000). It has been suggested that the RING domain is vital for the targeting of the E3 ubiquitination complex to proteins that have been marked for degradation (Borden, 2000). Although I have described the involvement of the RING finger domain in transcriptional repression and ubiquitination, it is likely that these domains are involved in far more cellular processes. However, the common theme that has been suggested for RING domain proteins is that they form large macromolecular scaffolds which enable proteins to be correctly positioned in biological processes (such as ubiquitination) and may also act as molecular modifiers, altering the specificity of these processes (Borden, 2000).

The B-box zinc finger is a 40 amino acid domain and is found in transcription factors, ribonucleoproteins and proto-oncoproteins but it does not have a clearly defined function (Borden, 1998). The B-box and RING finger motif in the RBCC/TRIM protein PML has been shown to be essential for the assembly and possible targeting of large macromolecular complexes, which I have already mentioned (Borden *et al.*, 1996). The coiled coil region is also important in forming high molecular weight complexes although the function of this domain has been shown to allow the self association of the Trim11 (Reymond *et al.*, 2001).

The B30.2 domain is a 160-170 amino acids domain, which is found in nuclear and cytoplasmic proteins, as well as transmembrane and secreted proteins, although the function of the domain is not known (Henry *et al.*, 1998). The SPRY domain is of unknown function but was named after the SP1a and Ryanodine receptors which mediate calcium release from the sarcoplasmic or endoplasmic reticulum (Ponting *et al.*, 1997).

Other members of the Trim protein family have been implicated in development, cell growth and identifying cell compartments although the specific function of Trim11 was speculated to be in the identification of cell compartments (Reymond *et al.*, 2001). However, it has recently been shown that Trim11 interacts with Humanin, which is a protein that was initially isolated from a cDNA library of an Alzheimer's disease patient's brain and was shown to stop neurotoxicity (Niikura *et al.*, 2003). Humanin has been shown to be expressed in neurons of the occipital lobe and glial cells of Alzheimer's disease brains but not in age matched controls (Tajima *et al.*, 2002). The interaction between Humanin and Trim11 is reliant on the coiled coil and B30.2 domains of Trim11 with the coiled coil domain instigating self association of Trim11 so that there is sufficient interaction of the B30.2 domain with Humanin (Niikura *et al.*, 2003).

It has also been shown that the expression of Trim11 negatively regulates Humanin expression, which is consistent with the function of the RING finger domain of Trim11 in protein ubiquitination and transcriptional repression (Niikura *et al.*, 2003). However, this report concludes that the regulation of Humanin levels is through ubiquitin-mediated protein degradation pathways due to the isolation of increased ubiquitinated Humanin protein in the presence of Trim11 (Niikura *et al.*, 2003).

Therefore, it is interesting to speculate that the interaction of Trim11 with Pax6 in regions of the brain may allow the specific regulation of Pax6 protein expression levels. Another related hypothesis is that Trim11 allows the formation of large macromolecular complexes and that the

interaction of Trim11 and Pax6 means that Pax6 is positioned correctly so that it can carry out its normal function. Disruption of these complexes due to mutation of either Pax6 or Trim11 may then lead to dysregulation of the protein components of the complex leading to human disease.

## 5.7 Expression Studies.

By looking at the UniGene website (see 2.1.6) I could quickly investigate the expression pattern of these genes. This revealed that all of the interactors have similar expression patterns to *PAX6*, with them all being represented by EST (Expressed sequence tagged) sequences from the eye and brain.

### 5.7.1 Expression of the interactor genes in the cDNA library.

As the genes encoding the potential interacting proteins, Homer3, Dncl1 and Trim11 had been isolated from the yeast 2-hybrid system several times each then a concern was that these genes were highly represented in the mouse brain cDNA library and therefore more likely to be pulled out in the library screens.

I decided to quantify the amount of DNA that was present in the cDNA library for each of the interacting proteins as well as *Pax6* and 2 controls: Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) gene, a housekeeping gene which has high expression (Patel *et al.*, 2002); and *Atp5a1*, the alpha subunit gene of ATP synthase, which has strong expression in the brain (Personal communication, B. Pickard). This was done by performing a PCR (see 2.4.2) with the correct PCR primers (see Table 2.4) at increasing number of cycles (20 to 35 cycles) to amplifying the relevant DNA bands from the mouse brain cDNA library. One tenth of the PCR product was then run out on a 0.8% agarose gel (see 2.2.7) as shown in Figure 5.9.

This showed that the *Pax6*, *Homer3*, *Dncl1* and *Trim11* genes were all equivalently expressed in the mouse brain cDNA library and showed lower expression than the two highly expressed controls. Therefore, this verified that the interacting proteins were not pulled out of the mouse brain cDNA library screens multiple times merely because they are more prevalent in the cDNA library.

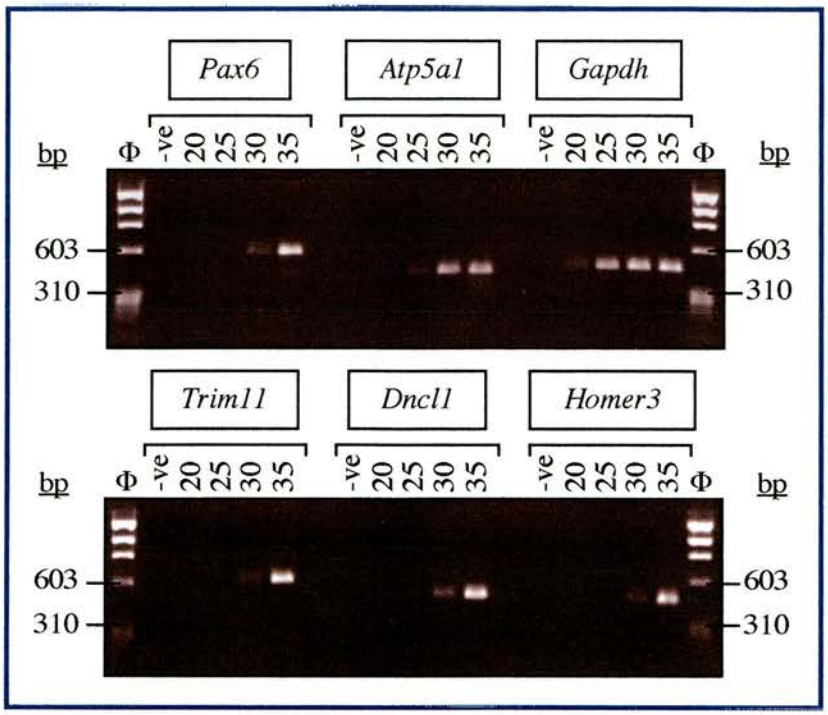



**Figure 5.9 Quantitation of *Pax6* and the three interactors in the cDNA library.**

Amount of cDNA that represents *Pax6* (600bp band) and the *Homer3* (485bp band), *Dncl1* (485bp band) and *Trim11* (609bp band) was estimated by performing PCR and varying the number of cycles from 20 to 35. Two controls, *Atp5a1* (415bp band) and *Gapdh* (450bp band), which represent highly expressed proteins in the brain, were also included. One tenth of the PCR was run on this 0.8% agarose gel along with Φx marker.

### 5.7.2 Expression of the interactor genes in relevant cell lines.

Although it was encouraging that the three interactor genes were expressed in similar regions as determined by looking at EST sequences, I wanted to specifically investigate whether these genes were expressed in a mouse and a human lens epithelium cell line (mv+ and H36CE2 respectively, see Table 2.8). These were known to express the PAX6 protein (Veronica van Heyningen, personal communication). The primers in Table 2.4 were designed against the exonic sequences of the mouse and human *Homer3*, *Dncl1*, *Trim11* and *PAX6* genes and these were used to performing RT-PCR (see 2.4.4) on the RNA (see 2.3.1) of the mouse (mv+) and human (H36CE2) *PAX6* expressing cell lines.

| Lens Epithelium Cell Line | PAX6 | Homer3 | Dncl1 | Trim11 |
|---|---|---|---|---|
| mv+ (Mouse) | +++ | +++ | +++ | + |
| H36CE2 (Human) | ++++ | ++ | +++ | + |

**Table 5.9 Expression of *PAX6* and interactors in lens epithelial cell lines.**

This table estimates the level of expression of *PAX6* and the three interactors (*Homer3*, *Dncl1* and *Trim11*) by RT-PCR in mouse (mv+) and human (H36CE2) lens epithelial cell lines.

This was very encouraging as the results of the RT-PCR (see Table 5.9) showed that *PAX6* and *Homer3*, *Dncl1* and *Trim11* are all expressed in the mouse and human lens epithelium cell lines. This experiment does not confirm that the proteins are expressed in these cells or that they interact *in vivo*, but the mRNA is expressed in the same cells which made me optimistic of a true interaction.

### 5.7.3 Sub-cellular localisation of PAX6.

The three novel and interesting interacting proteins that I have isolated are all predominately cytoplasmic in expression and it is known that PAX6 is mainly nuclear in its expression pattern. However, as discussed in section 1.3.3 there is a minus paired domain isoform of PAX6 (-PD PAX6), which has been shown to have a cytoplasmic localisation (Carriere *et al.*, 1995) and has also been shown to be highly expressed in the mouse brain by RT-PCR (Mishra *et al.*, 2002). Therefore, it is possible that the interacting proteins may associate with this isoform of PAX6 *in vivo*.

The subcellular localisation of PAX6 was investigated in mouse lens epithelium cells (mv+) using the subcellular proteome extraction kit (see 2.10.6). The resulting subcellular fractions were then run out on a 12% SDS-polyacrylamide gel (see 2.10.1), transferred to Hybond-P (see 2.10.2) and western blotting (see 2.10.3) carried out using the C terminal PAX6 antibody (see Table 2.9) and the HRP conjugated anti-rabbit secondary antibody (see Table 2.10). The results are shown in Figure 5.10.

It is interesting to discover that using this protocol the cytosolic protein fraction not only contains the –PD PAX6 isoform (24.3kDa) but also contains the –5a PAX6 isoform (46kDa), whereas only the +5a PAX6 isoform (48kDa) was located in the nucleic protein fraction. The results of this experiment confirm that the –PD PAX6 isoform is predominantly expressed in the

cytoplasm, which is encouraging for the potential interaction of PAX6 and the 3 novel interacting proteins.

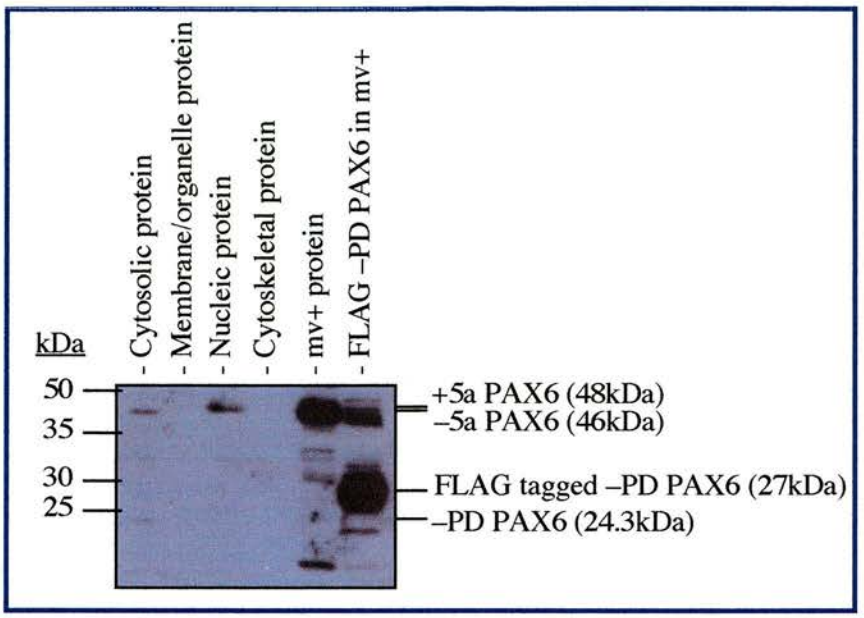



**Figure 5.10 Subcellular localisation of PAX6.**

This shows that there are 2 bands in the cytosolic protein fraction: -5a PAX6 at 46kDa; and –PD PAX6 at 24.3kDa. The nucleic protein fraction contains a single band representing +5a PAX6 at 48kDa. mv+ is protein extracted from mouse lens epithelial cells. The FLAG –PD PAX6 lane contains lysate isolated from mv+ cells that have been transfected with p3XFLAG-PD_PAX6 (see Table 8.3) that represents the 27kDa FLAG tagged –PD PAX6 produced. Autorad has been exposed for 2mins of an immunoblot probed with the C terminal PAX6 antibody (see Table 2.9) and the HRP conjugated anti-rabbit secondary antibody (see Table 2.10).

The fact that the +/- 5a PAX6 isoforms have different subcellular localisation with the +5a isoform being nuclear and the –5a isoform being cytoplasmic is also interesting. If this result is true then the –5a PAX6 isoform could potentially interact with these 3 interactors like the –PD PAX6 isoform.

It has been shown from experiments on quail Pax6 (Carriere *et al.*, 1995) that there are potentially two nuclear localisation signals (NLS) in the Pax6 protein. One of these, which is contained within exon 6 (see Figure 1.3), is thought to function in localising the paired domain containing Pax6 isoforms (i.e. +/- 5a Pax6) to the nucleus. The other NLS, which is thought to function in the –PD Pax6 isoforms, is located just before the homeodomain (LKRKLQR, aa 206-212, see Figure 3.1) and is similar to a SV40 nuclear localisation signal (Zhang *et al.*, 1998).

However, even these two NLS are not sufficient to direct non-nuclear proteins into the nucleus, and it is speculated that another region is required for correct nuclear localisation (Carriere *et al.*, 1995). It was shown by deletion studies that a region between amino acids 271-342 in Pax6 might also be involved in nuclear localisation, and this region shows similarity to a proposed bipartite nuclear signal (266-276) at the end of the homeodomain which might also have an involvement in nuclear localisation (Carriere *et al.*, 1995). However, even with the characterisation of these possible NLS it was still noted that a small proportion (7%) of transfected quail neuroretinal cells had a cytoplasmic localisation of the –5a Pax6 isoform (Carriere *et al.*, 1995), which suggests that there is potentially another mechanism that partitions Pax6 between the nucleus and the cytoplasm.

It has also been speculated that there might be a region within the PST domain that is required for cytoplasmic retention (Zhang *et al.*, 1998). As the PST domain contains several potential phosphorylation sites then it is possible that the subcellular localisation of Pax6 could be regulated by protein phosphorylation, which is already known to control nuclear transport (Jans and Hubner, 1996). Therefore, this is particularly interesting as I have already proposed the possibility of a link between MAPK (mitogen-activated protein kinase) and Pax6 due to the interaction between Pax6 and Homer3, and this could potentially be involved in the nuclear localisation of Pax6 isoforms.

However, it is also possible that the subcellular localisation of Pax6 could be modulated by the binding of other proteins that could potentially mask nuclear localisation signals. This has specifically been shown in cerebellar granule cells, as a protein called mDia binds to Pax6 in the nucleus (Tominaga *et al.*, 2002). This is sufficient to reduce the transcriptional activity of Pax6 by dissociating it from DNA and also alter its subcellular localisation making it more cytosolic (Tominaga *et al.*, 2002). Although not specifically characterised it is proposed that mDia binds across the Pax6 homeodomain, which is thought to contain a potential NLS, and it could potentially be this masking of the NLS that localises Pax6 to the cytoplasm.

Furthermore, recent work has determined that there is variability in the subcellular localisation of Pax6 even within a single cell type of a normal chicken retina (Shin *et al.*, 2003). It was shown that although Pax6 was nuclear in ganglion cells and the inner nuclear retinal layer, it was localised in the intercellular space and the cytoplasm in the outer nuclear retinal layer (Shin *et al.*, 2003). Therefore, this suggests that the subcellular localisation of the PAX6 protein is variable, and can fluctuate in different cell types (Shin *et al.*, 2003).

The +/- 5a isoforms of PAX6 are known to have unique functions during development (Mishra *et al.*, 2002), but the importance of the +5a Pax6 isoform in lens development is illustrated by the observation that transgenic mice that have a loss of function (Singh *et al.*, 2002) or an overexpression (Duncan *et al.*, 2000) of this isoform develop lens defects. As the +5a PAX6 isoform has been shown to be specifically important in lens development then it is very interesting that I found that this isoform is nuclear whereas the –5a PAX6 and –PD PAX6 isoforms are cytoplasmic. It is interesting to hypothesise that in this cell type the –5a and –PD PAX6 isoforms are either sequestered in the cytoplasm by a cytoplasmic retention factor or are unable to translocate to the nucleus due to masking of their potential NLS.

Although this is an intriguing result for lens epithelial cells, this subcellular localisation of Pax6 isoforms cannot be assumed for every cell type and it is possible that it could be different in brain neurons. Therefore, it is feasible that any of the PAX6 isoforms could be cytoplasmic before being translocated into the nucleus to elicit their individual gene expression pattern and that the selection of which isoform is re-located is cell type specific.

### 5.7.4 The next step.

Having provided more *in vivo* evidence that the expression of PAX6 and the three identified interactors overlap, another important question needed to be answered. "Could the interaction between these three interacting proteins and PAX6 be re-confirmed using the yeast 2-hybrid system in a pairwise manner?" This is discussed in chapter 6.

# Chapter 6 Pairwise Interaction

## 6.1 Introduction.

As described in chapter 5, the yeast 2-hybrid library screens using the C terminal and PST domain constructs (pDBLeu_TTST_norm and pDBLeu_ASNT_norm respectively) isolated three potentially novel and interesting interactors of PAX6 several times from the mouse brain cDNA library. However, to add weight to the argument that PAX6 does indeed interact with these 3 interactors, I needed to confirm these interactions in a pairwise manner i.e. PAX6 and each of the 3 interactors specifically introduced into yeast cells to investigate interactions.

As all the isolated *Homer3* clones were devoid of the first 72 amino acids of the protein (see 5.5.1), I also decided to clone the full length *Homer3* (see 6.2) and use this to further characterise the interaction with PAX6. I also decided to investigate the effect on these interactions with full length PAX6 (see Table 4.2) and to investigate the role of the C terminal domain of PAX6 on these interactions. Therefore, I screened the 3 interactors with the C terminal mutant forms of PAX6 (see 1.7.6 and Tables 4.3 and 4.4) and the minus C terminal domain PAX6 construct (see Table 4.5).

## 6.2 Cloning full length *Homer3* into pPC86.

The *Homer3* clone that was isolated from the yeast 2-hybrid screen was found to be missing the first 72 amino acids of the published IMAGE clone sequence. Therefore, a full length *Homer3* clone was created to investigate any differences with the truncated *Homer3* clone in its interaction with PAX6. A full length *Homer3* IMAGE clone was identified by BLAST searches (IMAGE: 3602414). The clone was ordered and DNA was prepared (see 2.5.3). The missing 70 amino acids were then PCR amplified from the IMAGE clone DNA using PCR primers that were designed with a 5' *Sal*I site and a 3' *Eco*RI site (see Table 2.3). The restriction sites used in this cloning strategy were chosen so that the PCR generated fragment could be easily added to the pPC86 *Homer3* clone using a *Sal*I site within the polylinker and a convenient *Eco*RI site near the 5' end of the pPC86 *Homer3* insert. The sequence verified PCR product was then cut with *Sal*I and *Eco*RI and ligated into the pPC86 *Homer3* clone that had been cut with the same restriction enzymes to generate the full length Homer3 clone (see Table 6.1).

| Construct Produced. | Description. | Schematic Diagram. |
|---|---|---|
| pPC86_Homer3_Full | Corresponds to the entire *Homer3* cDNA sequence. |   |

**Table 6.1 Full length *Homer3* in pPC86.**

The pPC86 vector produces a protein coupled to the GAL4 activation domain (purple oval). This table shows the artificially constructed full length *Homer3* construct that was created, which is also illustrated as a schematic diagram (Domains as in Table 5.8).

## 6.3 Investigating pairwise interactions.

Interactions were tested by first transforming yeast with a relevant pDBLeu PAX6 construct and one of the pPC86 interactor constructs (see 2.8.3). These pairwise interactions were then analysed using the full screen with an optimised 3AT concentration of 10mM 3AT that was sufficient to eliminate basal growth between interactors and PAX6 constructs.

### 6.3.1 Controls included in pairwise screens.

Apart from the transactivation ability of the whole PST domain constructs which will be discussed in Chapter 7, none of the remaining pDBLeu PAX6 or pPC86 interactor constructs had any transactivation potential when included in the yeast 2-hybrid system along with the corresponding empty vector, e.g. pPC86_Homer3 included with empty pDBLeu vector was not able to activate reporter gene expression. In addition to these controls the pPC86 interactor constructs were also tested against random pDBLeu constructs, which encoded SIX1, SIX3, Synip and Grg5. This showed that none of the interactor clones interacted with these non-specific proteins. PST domain and C terminal domain constructs containing the wild type alleles from patients REMAR and WHEDA (see Table 4.4) were also included as controls in these pairwise interactions. These wild type constructs contain part of the 3'UTR but are expected to form the normal C terminus of PAX6 and hence behave like normal PAX6 constructs. This was confirmed by the pairwise interactions full screens as these constructs did not behave differently from the normal PST or C terminal domain constructs (Data not shown).

### 6.3.2 Quantification of pairwise interaction results.

The full set of results obtained from this pairwise screening between pDBLeu PAX6 constructs and each of the pPC86 interactor constructs is illustrated in Table 6.2. By comparison to the 5 yeast control strains A to E which have increasing reporter gene expression (A, lowest and E, highest expression) I attempted to quantify the pairwise interactions so that the results could be interpreted more readily. By combining this level of reporter gene expression with the predicted colony phenotype, using Table 5.2, I quantified the interactions from ✓✓✓ as the strongest interaction to × which shows no interaction and no reporter gene expression. The interactions that have been termed 'weak' are those interactions where the reporter gene expression is very low when compared to the control yeast strains and would normally be characterised as a ✓ or ×, however, there is one inconsistent reporter gene phenotype.

It is clearly illustrated (see Table 6.2) that the whole PST domain has been confirmed to interact with Homer3 (full and truncated constructs), Trim11 and Dncl1. Although there is innate transactivation ability of the PST domain construct (pDBLeu_ASNT_norm), its interactions with Homer3, Dncl1 and Trim11 causes a higher amount of reporter gene expression when compared to the yeast control strains A to E than only the PST domain construct. Therefore, it can be concluded that the PST domain of PAX6 and these 3 interactors really do interact in this system and that it is not just an artifact of the transactivation ability of the PST domain.

The confirmation of the interactions between the C terminal domain PAX6 construct and Homer3 and Dncl1 was more problematic. In the full screen of *HIS3* positive colonies isolated in the C terminal domain library screen, these colonies showed low level *URA3* and *lacZ* reporter gene expression, indicating a weak interaction. However, when the pairwise interaction full screen was performed only the test for low level reporter gene expression was affected. This showed very slight inhibition of colony growth on plates containing 0.2% 5FOA (see 2.8.7), but there was no expression of the *lacZ* reporter gene. This result was not affected by repeating the full screen with freshly transformed yeast cells. I can only conclude that the reporter gene expression in these pairwise experiments is so low that it isn't sufficient to give a colour change in the *lacZ* assay, and only slightly affects the 5FOA result. Therefore, due to the inconsistency of the reporter gene expression I cannot confirm the pairwise interaction between the C terminal domain PAX6 construct and the pPC86 Homer3 or Dncl1 constructs with this system.

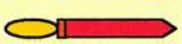
| Construct Name ↓ Schematic → | | pPC86 constructs: Homer3 | Homer3_Full | Dncl1 | Trim11 |
|---|---|---|---|---|---|
| **Normal pDBLeu constructs** | | | | | |
| PAX6_Full |  | × | × | × | × |
| ASNT_norm |  | ✓✓✓ | ✓✓ | ✓✓ | ✓✓ |
| SYSC_norm |  | ✓✓ | × | × | ✓✓ |
| MQTH_norm |  | Weak | × | × | × |
| TTST_norm |  | × | × | × | × |
| **Q422R mutant pDBLeu constructs** | | | | | |
| ASNT_mut |  | ✓✓ | ✓ | ✓ | ✓✓ |
| SYSC_mut |  | ✓ | × | × | ✓✓ |
| MQTH_mut |  | × | × | × | × |
| TTST_mut |  | × | × | × | × |
| **X423L (REMAR) mutant pDBLeu constructs** | | | | | |
| ASNT_RMmut |  | ✓ | × | Weak | ✓✓ |
| TTST_RMmut |  | × | × | × | × |
| **1615del10 (WHEDA) mutant pDBLeu constructs** | | | | | |
| ASNT_WHmut |  | Weak | × | × | ✓✓ |
| TTST_WHmut |  | × | × | × | × |
| **Minus C terminal pDBLeu construct** | | | | | |
| ASNT-TTST |  | Weak | × | × | ✓✓ |

**Table 6.2 Pairwise interaction results.**

This table shows the results of all the pairwise interactions between the various PAX6 constructs and the interactor constructs. A schematic diagram illustrates each construct. The interaction is quantified with ✓✓✓ being a definite interactor and × being a definite non-interactor and 'weak' indicating that there is low or no reporter gene expression, but that there is one inconsistent reporter gene phenotype. Schematic diagrams of pDBLeu PAX6 constructs as in Chapter 4 and pPC86 interactor constructs as in Tables 5.9 and 6.1.

Also of concern was the finding that full length PAX6 was shown not to interact with the 3 interacting proteins in this system. It could potentially be argued that as these 3 interactors do not interact with full length PAX6 then this weakens the proposal that the interactions occur *in*

*vivo*. However, in the yeast 2 hybrid system, the GAL4 DNA binding domain is coupled to the N terminus of the fusion protein when expressed, which could potentially alter the structure of the full length PAX6 protein. The GAL4 DNA binding domain is in fact roughly the same size as the paired domain (about 150 amino acids in length) and when coupled to the N terminus of the paired domain could sufficiently alter the conformation of PAX6 so that interaction with the 3 interactors is not possible. Furthermore, as the 3 interactors are predominantly cytoplasmic and a potential minus paired domain isoform of PAX6 (see 1.3.3) is also mainly cytoplasmic in expression, then interactions may occur due to the specific conformational arrangement of this –PD isoform. However, the characterisation of the –PD isoform and the interactors was not investigated using this yeast 2-hybrid system.

It is also interesting to note that although only the whole PST domain of PAX6 interacts with Dncl1, the interaction of Homer3 and Trim11 also occurs with the more truncated PAX6 construct pDBLeu_SYSC_norm, although this interaction is reduced with Homer3 and does not occur with full length Homer3. This potentially indicates that the interaction site of Homer3 and Trim11 with PAX6 may well occur more towards the C terminus of the PST domain than Dncl1. However, in the case of Homer3 this does not take into account the results of the C terminal mutant PAX6 construct interactions which will be discussed further in section 6.3.4.

### 6.3.3 Trim11 interaction site in PAX6.

Trim11 interacts specifically with the whole PST domain PAX6 construct and the more N terminally truncated construct pDBLeu_SYSC_norm. The PST domain constructs that either include the C terminal mutations (Q422R, X423L and 1615del10, see Figure 1.18) or lack the C terminal domain (see Table 4.5) do not alter this interaction (see Table 6.4 and Figure 6.1), which implies that the site of Trim11 interaction with PAX6 may be tentatively suggested. The interaction between Trim11 and PAX6 is not altered by the removal of the N terminal 75 amino acids of the PST domain (difference between the ASNT and SYSC constructs, see Figure 3.1) but there is no interaction between Trim11 and the pDBLeu_MQTH_norm construct. Therefore, I speculate that Trim11 interacts with PAX6 in the region illustrated in Table 6.3 around the SYSC motif of the PST domain (green box, see Figure 3.1). The amino acid sequence of PAX6 immediately surrounding this SYSC motif shows weak homology over evolution and may suggest a specific region for interactions. Mutational studies would be interesting to further characterise the interaction site between PAX6 and Trim11 in this region.

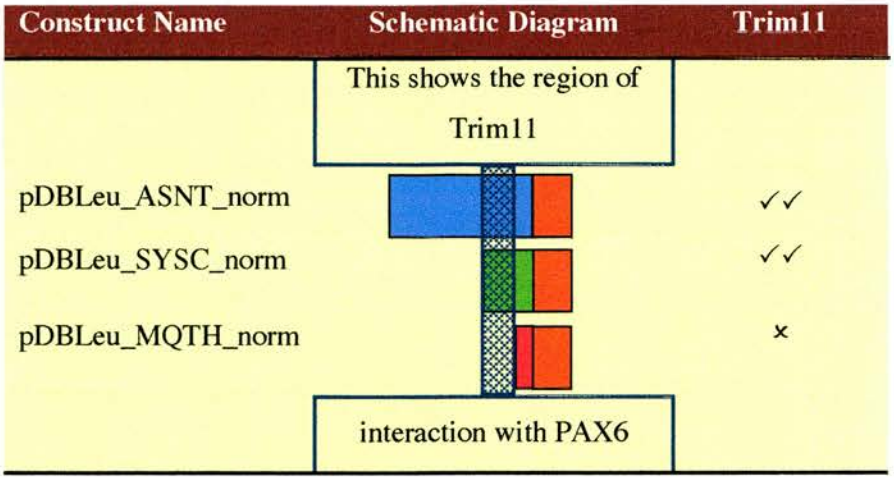



**Table 6.3 Illustration of possible Trim11 interaction site in PAX6.**

This table highlights the possibility that Trim11 may bind to PAX6 in a specific region indicated by the vertical stippled box. This is based on the different interacting potentials of the truncated PAX6 constructs. The GAL4 DNA binding domains (yellow oval) in previous schematic diagrams have been removed so that the interacting site can be clearly illustrated. Other domains of the schematic diagrams as in Table 4.1.

### 6.3.4 C terminal influence on pairwise interactions.

As already described, the interaction between the three proteins Homer3, Dncl1 and Trim11 and the PST domain of PAX6 was confirmed but the interaction with the C terminal domain was not. These results are consistent with Trim11 only being isolated from the PST domain library screen but Homer3 and Dncl1 were isolated from both library screens and I was concerned that I could not re-confirm their interaction with the C terminal domain construct. Although I have already mentioned that the interaction between Homer3 and Dncl1 and the C terminal domain of PAX6 could be extremely weak (see 6.3.2), I was reassured when I discovered that the C terminal domain plays an important role in these interactions (see Table 6.4).

Although there is an interaction between the PST domain of PAX6 and Homer3 or Dncl1, this interaction is severely affected by the removal of the C terminal domain (see Table 6.4 and Figure 6.1). In the case of full length Homer3 and Dncl1 the interaction with PAX6 seems to be dependent on the C terminal domain even though the interaction with the C terminal domain itself is very weak. Therefore, it is feasible to speculate that the interaction of the C terminal domain and the PST domain may generate a specific conformation that cannot be achieved without the intact C terminal domain.

Even more tantalising are the results that the interaction between Homer3 or Dncl1 and PAX6 is reduced or abolished due to C terminal mutations that are known to cause eye malformations in patients (see Table 6.4).

| PAX6 clone used in pairwise interaction. | | Homer3 | Dncl1 | Trim11 |
|---|---|---|---|---|
| **PST** | | ✓✓✓ | ✓✓ | ✓✓ |
| **Q422R Mutant** | | ✓✓ | ✓ | ✓✓ |
| **X423L Mutant** | RM | ✓ | Weak | ✓✓ |
| **1615del10 Mutant** | WHEDA | Weak | × | ✓✓ |
| **Minus C ter** | | Weak | × | ✓✓ |

**Table 6.4 Pairwise interaction results with various PAX6 mutants.**

This table shows the affect that PST domain constructs including the PAX6 C terminal mutants (see 1.7.6) or lacking a C terminal domain have on the interaction between PAX6 and the 3 isolated interactors Homer3, Dncl1 and Trim11. Schematic diagrams as in Chapter 4.

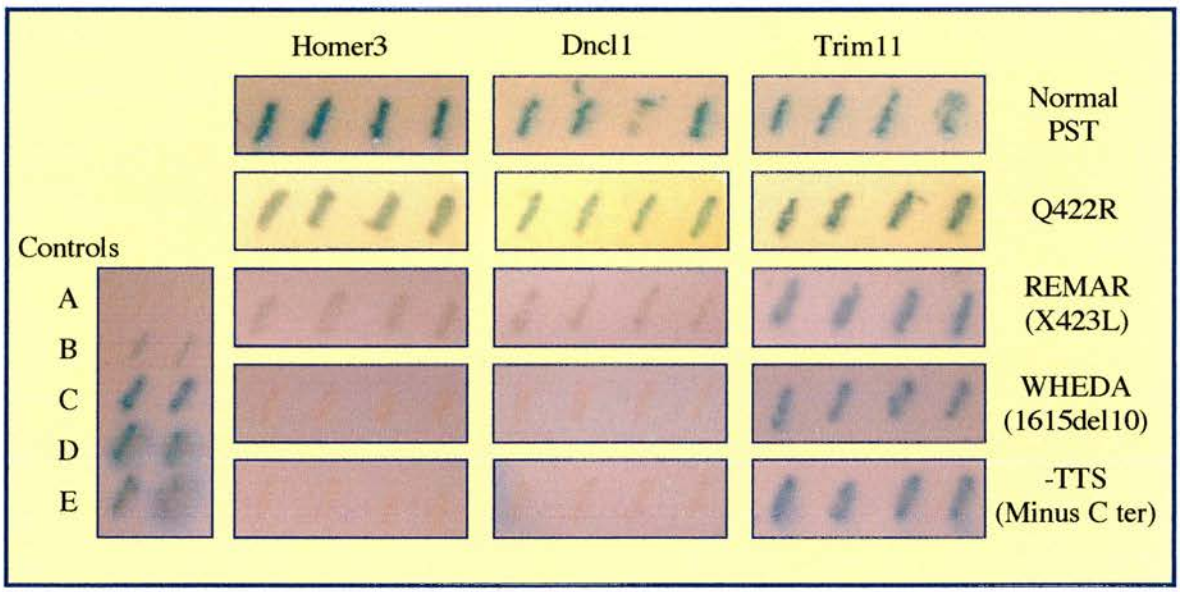



**Figure 6.1 Alteration of *lacZ* expression due to altered C terminus of PAX6.**

This figure shows how the interaction of Homer3, Dncl1, Trim11 and PAX6 is altered due to changes at the C terminus of PAX6, as shown by changes in *lacZ* reporter gene expression. The alterations of the C terminus of PAX6 are the 3 C terminal mutations (see 1.7.6) and a PST domain devoid of the C terminal domain (see Table 6.4 for schematic diagrams). For comparison, a control panel is included showing the yeast control strains A-E and their increasing expression of the *lacZ* reporter gene.

It has been shown for both Homer3 (full and truncated) and Dncl1 that the various C terminal mutations have an increasingly detrimental effect on the interactions: the Q422R mutation has the mildest effect and reduces the interaction with both Homer3 (full and truncated) and Dncl1; the X423L (REMAR) mutation reduces the interactions with both Homer3 and Dncl1 even further, abolishing the interaction with full length Homer3; and the 1615del10 (WHEDA) mutation has the most severe effect with the interaction with Homer3 (full and truncated) and Dncl1 being abolished (see Table 6.2 and Figure 6.1).

## 6.4 Discussion.

The Q422R mutation changes the very last amino acid of PAX6 and is found in patients showing either mild aniridia characterised by uveal ectropion (Azuma and Yamada, 1998) or more classical aniridia with foveal hypoplasia (Singh *et al.*, 2001).

The X423L (REMAR) mutation occurs in several patients all showing aniridia. The mutation is an insertion of a T nucleotide into the normal stop codon of *PAX6*. The amino acid sequence of the PAX6 protein up to the normal stop codon is not altered but the mutated protein encodes a further 35 amino acid extension into the 3'UTR (untranslated region). This mutation has also been associated with the incorrect formation of brain structures such as the anterior commisure and neuronal connections to the olfactory bulb. In one case the mutation was also linked with brain anomalies that were manifested by autistic type behaviour (Chao *et al.*, 2003).

The 1615del10 (WHEDA) mutation is found in a family suffering from classical aniridia as well as an atypical neurobehavioural phenotype, namely impaired social understanding and verbal inhibition (Heyman *et al.*, 1999). This mutation causes the deletion of 10 nucleotides just before the normal stop codon of PAX6 and so the final 5 amino acids of the normal PAX6 protein are altered and the protein is elongated by a further 103 amino acid extension into the 3'UTR.

It is interesting to see that as more of the C terminal domain of PAX6 is altered by the various C terminal mutations (see Figure 1.18) then the interaction with Homer3 or Dncl1 decreases (see Table 6.4 and Figure 6.1). I have already suggested that a potential function for these proteins is in neurogenesis, hypothesising that PAX6 could be sequestered in the post-synaptic density due to its interaction with Homer3 and upon activation of the post synaptic density by extracellular stimuli, MAPK activated PAX6 could then be transported to the cell nucleus (due to its Dncl1 interaction) to alter cell gene expression (see 5.6.3).

The finding that the mutations that alter the largest amount of the C terminal domain of PAX6 (1615del10, WHEDA) cause a more severe brain phenotype is intriguing. It is possible that the neurobehavioural phenotype of these patients may be due to incorrect neuronal connections in specific regions of the brain that may in part be caused by the inability of PAX6 to interact with HOMER3 and/or DNCL1. In a patient with the X423L mutation the defect in olfactory function was suggested to be due to incorrect neuronal patterning between the brain and the intact olfactory bulb (Sisodiya et al., 2001). More intriguing for this mutation is the discovery that 2 patients with this mutation both developed polymicrogyria, which is a rare brain defect that is thought to be caused by disruption of neuronal migration (Mitchell et al., 2003). Furthermore, it was noted that one of these patients also suffered from epilepsy, which may also allude to incorrect neuronal patterning of the brain, which made this patient more susceptible to seizures.

Conversely, it is interesting to note that the phenotype produced by the Q422R mutation, which only affects the final amino acid of the C terminal domain has not been reported to cuase any brain or behavioural abnormalities. This may be simply due to the fact that the brain phenotype was not investigated or it could be that as this mutation does not abolish interactions between PAX6 and HOMER3 or DNCL1, then normal brain development can occur.

It has already been suggested that the expression of PAX6 in the brain is required to provide local navigational information to developing axons (Mastick et al., 1997). Therefore, these PAX6 mutations could alter or disrupt the normal gradient of PAX6 that occurs in regions of the developing brain (see 1.5.2) resulting in the associated brain abnormalities of these aniridia patients.

However, it is interesting to note that Homer mutants result in axon pathfinding errors in *Xenopus* optic tectal neurons (Foa et al., 2001) suggesting that there may be a possible link between these scaffolding proteins and proteins that govern axon trajectories. Although my suggestion of the function of the interactions between PAX6 and Homer3 and Dncl1 is in axonal trafficking from the dendrites to the cell nucleus, it is possible that these interactions could be involved in axon guidance due to the activation of genes that may be involved in axonal guidance during brain development (Meech et al., 1999).

Therefore, the reduced or disrupted interaction of the mutant PAX6 and these interacting proteins perhaps suggests that as the mutation of the C terminus of PAX6 becomes more severe then either PAX6 is not sequestered in the post-synaptic density by Homer3 or is not able to be

transported to the nucleus by Dncl1 to alter gene expression. This may mean that the gradient of PAX6 is not set up correctly causing changes in axon guidance or the altered gene expression may cause problems in neuronal patterning that are evident in patients that contain these PAX6 mutations. However, at this stage this is only a hypothesis attempting to link my findings with the phenotype of patients with these PAX6 mutations, and it must be remembered that these phenotypes may have an alternative explanation.

# Chapter 7 Transactivation

## 7.1 Introduction.

It was noticed that when the PST domain *HIS3* positive colonies were screened using the X-gal assay (see 5.2), the majority of these colonies as well as the pDBLeu_ASNT_norm only control had a positive blue colour. This indicated that the innate transactivation ability of the entire PST domain (Tang *et al.*, 1998; Mikkola *et al.*, 1999) was functioning in the yeast 2-hybrid system and was capable of activating reporter gene expression.

Therefore, this transactivation ability of the PST domain was investigated further and the transactivation ability of constructs that contained the various C terminal mutations was examined.

## 7.2 Yeast 2-hybrid reporter gene expression.

Figure 5.2 shows an illustration of how the yeast 2-hybrid system works to activate reporter gene expression. The PST domain pDBLeu construct produces a GAL4 DNA binding domain coupled protein that will bind to the DNA binding site and the empty pPC86 vector will produce the uncoupled GAL4 activation domain. However, the first question that I wanted to answer was "Is the transactivation of the reporter genes by the PST domain dependant on the GAL4 activation domain in this system?"

### 7.2.1 PST domain and the GAL4 activation domain.

I answered this question by transforming the PST domain construct (pDBLeu_ASNT_norm) into yeast cells with and without the empty pPC86 vector that encodes the GAL4 activation domain. I also did the same for the C terminal domain construct (pDBLeu_TTST_norm) to confirm that this construct has no transactivation ability. It must be remembered that the transformed yeast cells lacking pPC86 do not have the ability to synthesise tryptophan since pPC86 carries the *TRP1* gene. Therefore, the plates used in the reporter gene assays (see 2.8.7) had to include tryptophan to allow growth of these yeast cells. The full range of reporter gene assays were then performed as normal and the results are represented in Figure 7.1.

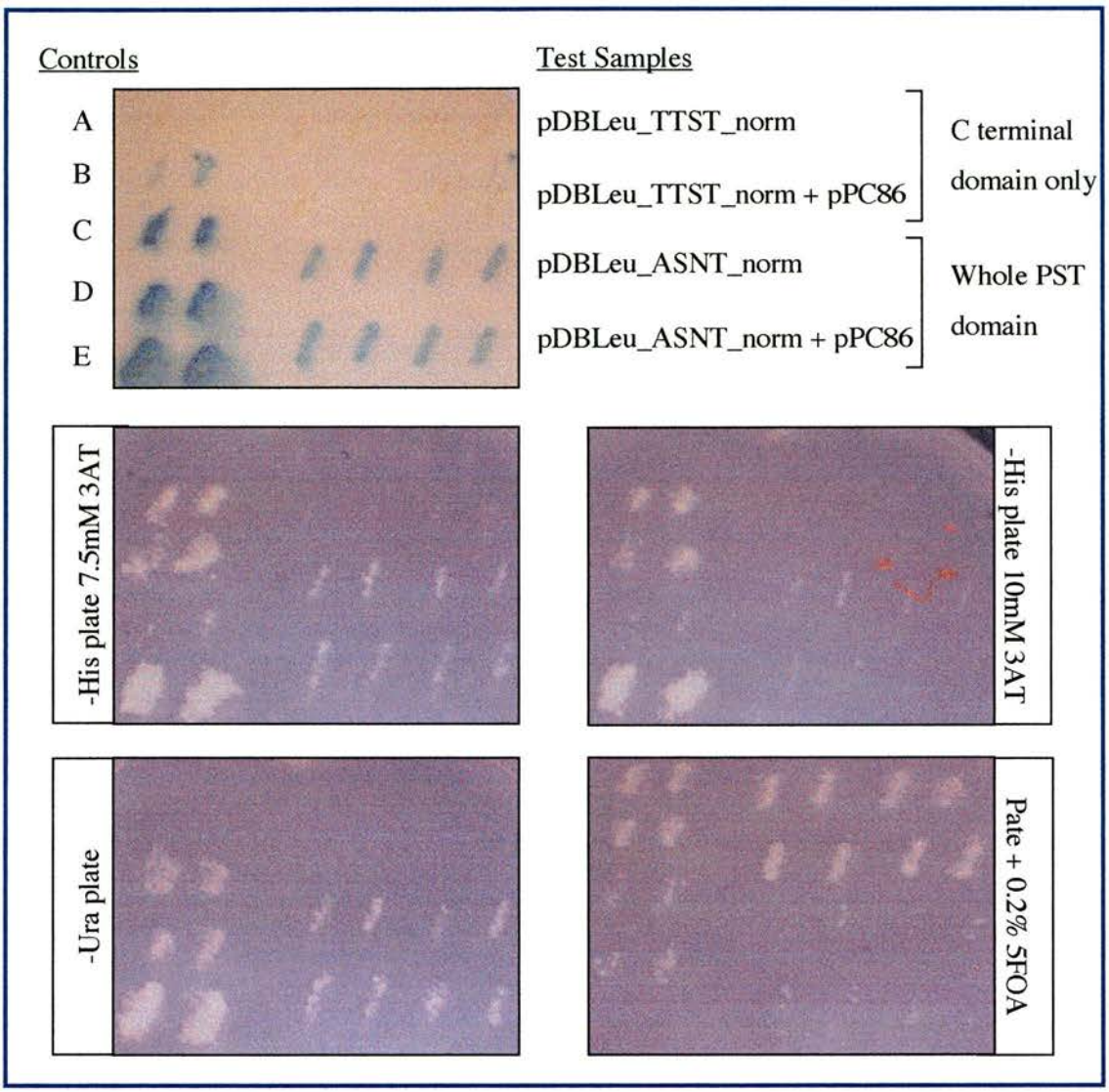



**Figure 7.1 Assessing transactivation ability without the GAL4 activation domain.**

This figure shows the results of all the reporter gene assays. Each panel illustrates the various plates containing selective media and each plate has the same layout as illustrated for the X-gal assay plate. Two concentrations of 3AT were used: 7.5mM for the C terminal construct and 10mM for the PST domain constructs. For more details about this full screen see 2.8.7.

This confirmed that the C terminal domain construct was unable to activate the reporter genes, even at low levels (tested by inhibition of growth on 5FOA containing plates) irrespective of whether the GAL4 activation domain was present or absent. However, it was clearly shown that there is transactivation of all the reporter genes with the PST domain construct and this expression is identical with and without the GAL4 activation domain. I also tested the more truncated PAX6 constructs (see Table 4.1) as well as full length PAX6 (see Table 4.2) as part of

the pairwise interaction experiments (see 6.3) and all these constructs had no transactivation ability. Therefore, only the PST domain construct has been shown to have transactivation ability and this is independent of the GAL4 activation domain. Initially this caused some concern, as I was worried that the pairwise results with the PST domain might result from its transactivation ability. However, as I've already mentioned in Chapter 6 (see 6.3.2) the results of the pairwise interactions with Homer3, Dncl1 and Trim11 all produced higher reporter gene expression than a PST domain only control when compared to the yeast control strains A to E.

### 7.2.2 Effect of C terminal mutations on transactivation.

Having ascertained that only the PST domain construct was capable of activating the reporter genes in the yeast 2-hybrid system, I wanted to investigate what effect altering the C terminus of the PST domain would have on this transactivation ability. Therefore, I transformed the pDBLeu PAX6 constructs that were lacking the C terminal domain (see Table 4.5) or that contained the C terminal mutations (see Tables 4.3 and 4.4) with the empty pPC86 vector into yeast cells. I then ran a full screen on the resultant transformed colonies and the interesting results are shown in Table 7.1.

| PAX6 clone used in pairwise interaction with empty pPC86. | Transactivation ability. |
|---|---|
| **PST** | ✓ |
| **Q422R Mutant** | Weak |
| **X423L Mutant** (RM) | × |
| **1615del10 Mutant** (WHEDA) | × |
| **Minus C ter** | × |

**Table 7.1 Varying transactivation of PST domain constructs.**

This table shows the varying transactivation ability of the PST domain constructs including the PAX6 C terminal mutants (see 1.7.6) or lacking a C terminal domain (see Table 4.5). Schematic diagrams as in Chapter 4.

I also included PST domain constructs containing the wild type alleles from patients REMAR (X423L) and WHEDA (1615del10, see Table 4.4) as controls. These constructs should produce normal PST domain proteins that terminate at the normal stop codon of *PAX6*. The results of the reporter gene assays showed that these wild type constructs behaved exactly the same as the normal PST domain construct (see Table 7.1) and activated the reporter genes.

Although the PST domain construct was shown to have transactivation ability it was very interesting to see that the constructs encoding the C terminal mutations caused a change in this transactivation, as is illustrated by the X-gal assay results in Figure 7.2.

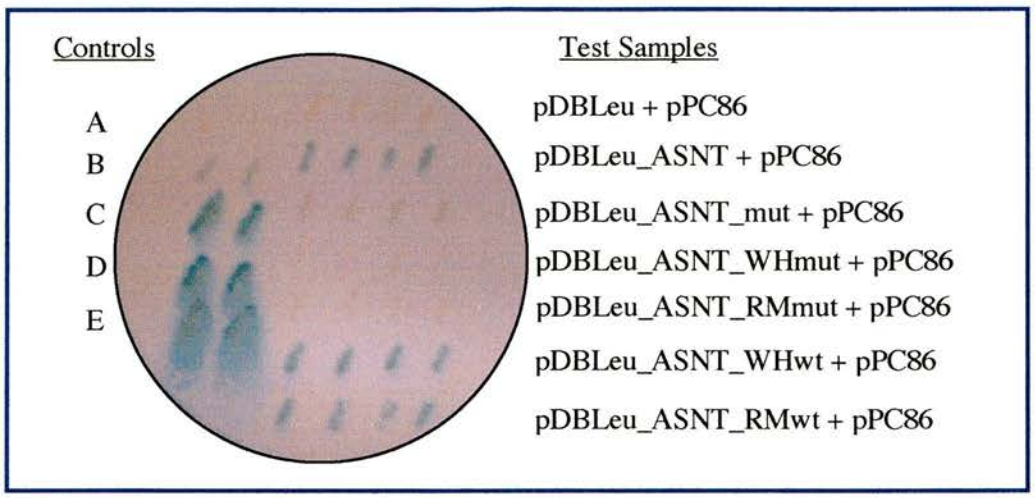



**Figure 7.2 PST domain transactivation of *lacZ* reporter gene.**

This shows how the transactivation ability of the PST domain is altered due to the various C terminal PAX6 mutations. X-gal assay on 7 test samples in quadruplicate and A-E control yeast strains in duplicate showing varying amounts of *lacZ* reporter gene expression. Construct names as in Chapter 4: ASNT refers to the PST domain; ASNT_mut contains the Q422R mutation; WHmut and WHwt are constructs that contain the WHEDA (1615del10) mutant and wild type alleles respectively; likewise RMmut and RMwt are constructs that contain the REMAR (X423L) mutant and wild type allele respectively.

The Q422R mutation changes just the final amino acid of the normal PAX6 protein; however, the inclusion of this mutation in a PST domain construct (pDBLeu_ASNT_mut) was sufficient to substantially reduce reporter gene activation. The transactivation ability of this construct was sufficient to cause inhibition of growth in the 5FOA assay (a test for low levels of *URA3* expression), but was insufficient to allow growth on plates lacking uracil (a test for high levels of *URA3* expression).

Although the Q422R mutant construct still retained a very weak transactivation ability, the PST domain constructs that contained the run on mutations 1615del10 (WHEDA) and X423L (REMAR) and produce elongated PAX6 proteins, had absolutely no transactivation ability. The PST domain construct that was lacking the C terminal domain was also sufficient to stop activation of reporter genes (see Table 7.1).

As I mentioned in the previous section I was concerned that the pairwise interaction results (see 6.3.2) between the PST domain PAX6 constructs and the interacting proteins (Homer3, Dncl1 and Trim11) could be an artefact of the transactivation ability of the PST domain. Furthermore, the validity of the pairwise interactions between the mutated PST domain constructs and Homer3 or Dncl1 could also be questioned, as the reduced interactions due to these mutations could feasibly be accounted by the reduction in the transactivation ability of the mutant constructs themselves. However, this theory can be discounted because even though the PST domain constructs with an altered C terminal domain do not show any transactivation ability, the interaction between these constructs and Trim11 is the same as with the normal PST domain construct. This therefore, goes some way in validating the interactions between the various PAX6 constructs and the three interactors and subsequent loss of interactions in some cases due to these C terminal mutations.

## 7.3 Transactivation due to a specific conformation?

Previous investigation into the functioning of the transactivation domain of PAX6 has shown that the entire PST domain is required to produce its transactivation ability (Tang *et al.*, 1998; Mikkola *et al.*, 1999). More specifically it has been shown that the C terminal 79 amino acids of PAX6 has a potent transactivation function (Czerny and Busslinger, 1995) and that the cooperation between the N (amino acids 271-331) and C (amino acids 380-416) terminal regions of the PST domain are needed for it to carry out its function as a transcriptional activator (Carriere *et al.*, 1995).

My results have substantiated these findings and demonstrate that the C terminal domain of PAX6 is vital in maintaining the transactivation ability of the PST domain. It is possible that the cooperation of the N and C terminal regions of the PST domain produce a specific conformation that instigates the transactivation potential. This would explain why there is no transactivation with only the C terminal domain of PAX6 or with the PST domain lacking the C terminal domain. A similar explanation was proposed by Mikkola *et al.*, (1999) who suggested the large size of the PST domain provided a potential interaction surface for protein interactors involved in the transcriptional machinery. Therefore, the specific conformation set up by the interaction of the N and C terminal regions of the PST domain could potentially be a site for the interaction of other proteins. Only when the specific conformation produces this correct interface can PAX6 interact with the correct set of proteins to carry out its function.

However, the most interesting results are that specific C terminal mutations that have been identified in patients with eye disease and some related brain anomalies (see 1.7.6) are sufficient to reduce or abolish the transactivation ability of the PST domain. Fitting this into the hypothesis that the PST domain forms a specific conformation suggests that these mutations alter this conformation. In the case of the Q422R mutation perhaps the effect of this mutation is less severe on the conformation of the mutated PAX6 as a limited amount of transactivation is still possible. However, the elongated proteins of 1615del10 and X423L are sufficient to completely alter the conformation of the PST domain and abolish transactivation ability. Therefore, if the proteins containing this missense mutation and these run on mutations are produced then the alteration to the conformation of the PST domain of PAX6 will mean that PAX6 cannot elicit its normal function.

# Chapter 8 Immunoprecipitation

## 8.1 Introduction.

Three novel and interesting protein interactors of PAX6 have been identified, Homer3, Dncl1 and Trim11. These interactions have been confirmed using the yeast 2-hybrid system in a pairwise manner (see Table 6.2) and it has been shown that these interactions are specifically affected by mutating the C terminal domain of PAX6 (see Table 6.4).

The next step of my PhD was to try and confirm these interactions using another method other than the yeast 2-hybrid system. To increase the validity of these interactions my aim was to confirm the interactions by *in vivo* or *in vitro* co-immunoprecipitation.

## 8.2 *In vivo* co-immunoprecipitation.

*In vivo* co-immunoprecipitation is a purification technique, which is used to determine whether two proteins are able to interact in cells. First the GST tagged interactor protein is expressed in mouse lens epithelium cells (mv+, see 2.9.5) and then a specific anti-GST antibody is added to the cell lysate (see 2.11.1). The antibody-protein complex is then pelleted using protein G sepharose beads (see 2.11.2), which binds to the GST antibody. If native PAX6 binds to the GST tagged interactor proteins then this will also be pelleted due to the interaction. Then identification of the proteins in the pellet can be determined by SDS-PAGE (see 2.10.1) and western blot (see 2.10.3).

Failing this, my next idea was to use specifically tagged interactor protein (GST tagged) and PAX6 protein (FLAG tagged –PD isoform) to see if this could be recovered using a similar co-immunoprecipitation technique.

## 8.3 Cloning interactors into pDEST27.

The GATEWAY cloning system (Invitrogen) was used to make constructs that would express Homer3, Trim11 and Dncl1 coupled to an N-terminal GST molecule. These constructs were created in an attempt to confirm the interaction with the PAX6 protein *in vivo*.

DNA of the appropriate pPC86 clone was cut with *Sal*I and *Not*I, and the cDNA insert was sub-cloned into the GATEWAY entry vector pENTR3c (also cut with *Sal*I and *Not*I). The LR

Recombination Reaction was then used to transfer the appropriate pPC86 derived fragments from the pENTR3c clones and into a GATEWAY destination vector pDEST27 (see 2.5.1).

A negative control was also generated by removing the negative selection *ccd*B gene from the pENTR4 GATEWAY entry vector with an *Eco*RI digest. The pENTR4 vector was used to generate this construct instead of the pENTR3c entry vector due to technical difficulties that meant that the *ccd*B gene could not be removed from the pENTR3c vector. The pENTR4 vector has a different polylinker region; otherwise it is identical to the pENTR3c vector. The LR Recombination Reaction was then used to transfer the empty cassette from the pENTR4 clone into the pDEST27 vector to generate the pDEST27-ccdb construct (see Table 8.1). Therefore, this construct will encode an N terminal GST tag and a stretch of 21 non-specific amino acids.

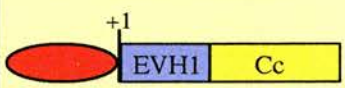
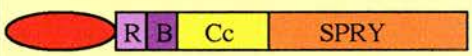
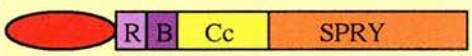
| Construct Produced. | Description. | Schematic Diagrams. |
|---|---|---|
| pDEST27_Homer3_T2 | Normal *Homer3* isolated clone is 285 amino acids. |   |
| pDEST27_Homer3_T24 | Additional N terminal amino acid. |   |
| pDEST27_Homer3_Full | Full length Homer3 with N terminal 70 amino acids. |   |
| pDEST27_Trim11_XA3 | Longest isolated *Trim11* clone. |   |
| pDEST27_Trim11_A47 | Commonly used *Trim11* clone. |   |
| pDEST27_Dncl1_XT31 | Longest isolated *Dncl1* clone. |  |
| pDEST27_Dncl1_A62 | A splice variant *Dncl1* clone that was isolated. |  |
| pDEST27-ccdb | Negative control only expressing GST tag. |  |

**Table 8.1 pDEST27 constructs.**

The pDEST27 vector produces N terminal Glutathione S-transferase (GST, red oval) fusion proteins, which are capable of high-level expression in mammalian cells. This table shows the various pDEST27 constructs that were created.

When the *Homer3* constructs, which are illustrated in Table 8.1, are expressed *in vivo* the full length and two truncated Homer3 proteins will be attached to an N terminal GST tag via a linker region of 12 amino acids, which is produced from the pDEST27 vector. However, the situation for the *Dncl1* and *Trim11* constructs (see Table 8.1) is different because these cDNAs contain 5'UTR as well as the protein coding region. In both cases the 5'UTR did not contain an in frame stop codon and so the GST tagged constructs will therefore, have an elongated linker region (see Table 8.2) between the GST tag and the encoded protein. However, it must be remembered that these linker regions would also have been present in the yeast 2-hybrid system attaching the Dncl1 and Trim11 proteins to the GAL4 activation domain. Therefore, the elongated linker regions between these proteins and the GST tag do not interrupt their interaction with PAX6.

| Construct | GST tag | Linker | Coded Protein | Total | kDa Size |
|---|---|---|---|---|---|
| |  | | | | |
| pDEST27_Homer3_T2 | 224aa | 12aa | 285aa | 521aa | 57.3kDa |
| pDEST27_Homer3_T24 | 224aa | 12aa | 286aa | 522aa | 57.4kDa |
| pDEST27_Homer3_Full | 224aa | 12aa | 357aa | 593aa | 65.2kDa |
| pDEST27_Trim11_XA3 | 224aa | 59aa | 467aa | 750aa | 82.5kDa |
| pDEST27_Trim11_A47 | 224aa | 42aa | 467aa | 733aa | 80.6kDa |
| pDEST27_Dncl1_XT31 | 224aa | 44aa | 89aa | 357aa | 39.3kDa |
| pDEST27_Dncl1_A62 | 224aa | 46aa | 89aa | 359aa | 39.5kDa |

**Table 8.2 Sizes of the expressed interactor protein constructs.**

This table shows the sizes in amino acids (aa) and kDa of the expressed GST tagged interactor protein constructs (see Table 8.1). The table breaks down the expressed proteins into the GST tag, a linker region and the encoded interactor protein and gives sizes in amino acids (aa) for each.

## 8.4 Cloning minus paired domain *PAX6* into p3XFLAG.

In mouse and quail a PAX6 isoform has been identified that is reported to be cytoplasmic and is lacking the paired domain (see 1.3.3). This could explain how PAX6 interacts with the 3 isolated interactors, which are all predominantly cytoplasmic in their expression patterns. The start of the minus paired domain (-PD) protein is in exon 8 of PAX6 at the amino acid sequence MRL (M202, see Figure 3.1).

The start of the –PD protein was created by PCR so that when inserted into p3XFLAG it will form an inframe protein with an N terminal FLAG tag. The EE4 PAX6 cDNA construct was used as the template DNA and the –PD fragment was PCR'd producing a fragment that extends from an artificially inserted 5' *Not*I site and up to the *Nco*I site in exon 11 of *PAX6* (see Figure 1.3).

The remainder of the cloning strategy included many subcloning steps. Firstly, the entire *PAX6* cDNA coding region was cut out of the EE4 construct with *Eco*RI and cloned into pBluescript II SK+ to create pBluescipt_EE4-RI. The pBluescipt_EE4-RI construct was then cut with *Not*I and *Nco*I to drop out the 5' part of the coding region including the paired domain, which was then replaced with the confirmed –PD PCR fragment. The final step included cutting the entire fragment that corresponds to the –PD *PAX6* clone and transferring it from pBluescript into p3XFLAG-CMV™-7.1 (Sigma) cut with *Not*I and *Eco*RI.

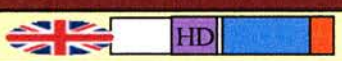
| Construct Produced. | Description. | Schematic Diagram. |
|---|---|---|
| p3XFLAG-PD_PAX6 | Minus paired domain isoform of PAX6. |   |

**Table 8.3 Minus paired domain PAX6 isoform.**

The p3XFLAG-7.1 expression vector produces N terminal FLAG fusion proteins, which are capable of transient intracellular expression in mammalian cells. This table shows the -PD *PAX6* p3XFLAG-7.1 construct that was created, illustrated by a schematic diagram.

## 8.5 Controls for *in vivo* co-immunoprecipitation.

So that the results of the *in vivo* co-immunoprecipitation experiments (see 8.6) could be accurately determined it was necessary that I include sufficient controls, which I will discuss in this section.

An aliquot of untransfected mv+ cell lysate was always included to ascertain any cross reactivity with the western blot antibodies. Also a sample of cell lysate from cells transfected with the relevant pDEST27 construct was always run alongside the co-immunoprecipitation reaction to make sure that transfection had occurred. In addition I transfected cells with the pDEST27-ccdb construct (see Table 8.1) and used the resulting cell lysate to check that the expected 28kDa protein was being expressed.

To check that the electrophoretic transfer (see 2.10.2) and western blotting (see 2.10.3) was adequate to transfer and detect the relevant proteins, I always included a GST or a PAX6 control where applicable. As a GST control I used a commercially available 26kDa GST protein (Sigma) and for a PAX6 control I used an aliquot of purified mv+ protein (see 2.10.4).

Two negative controls were also always included when performing the co-immunoprecipitation reactions. A co-immunoprecipitation reaction was carried out using the cell lysate of cells transfected with the pDEST27-ccdb construct, which should purify only the N terminal GST tag. I also included a co-immunoprecipitation reaction that used no cell lysate to determine any non-specific banding caused by the co-immunoprecipitation protocol.

## 8.6 Results using the GST antibody.

The first thing that I needed to confirm was that the GST tagged interactor proteins (see Table 8.1) were being expressed in the mouse lens epithelium cell line (mv+). Therefore, the DNA of each of the interactor constructs was individually transfected (see 2.9.5) into this cell line and the resulting cell lysates analysed by SDS-PAGE (2.10.1) and western blot (2.10.3).

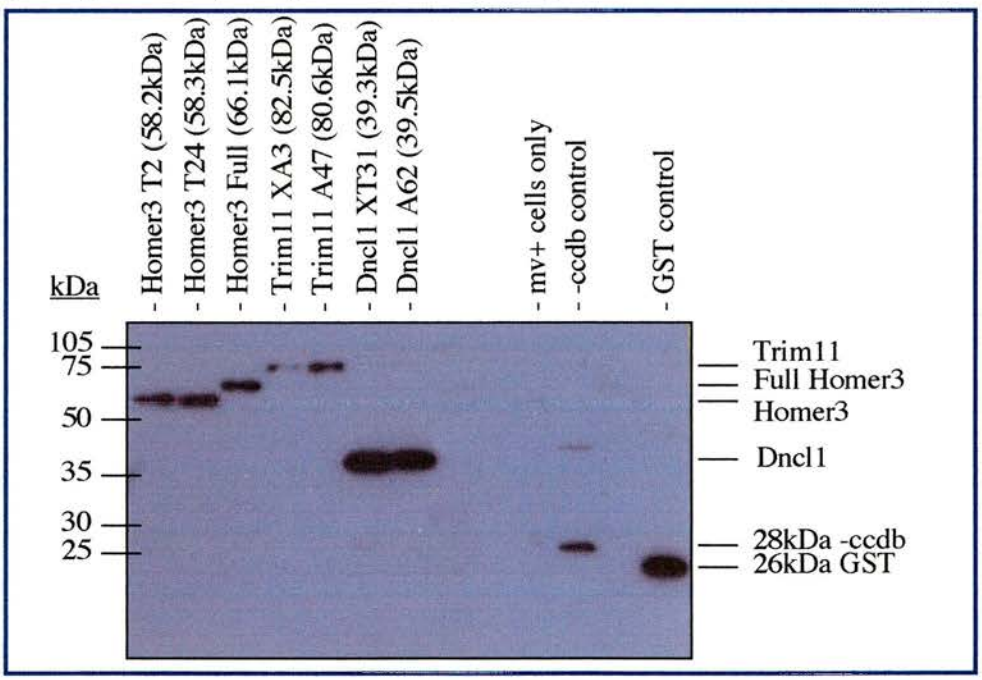



**Figure 8.1 Transfection and expression of GST tagged interactor constructs *in vivo*.**

This shows the resulting bands after 2.5μg of DNA transfected into mv+ cells. 10μl samples of lysate run on a 12% acrylamide gel. Autorad has been exposed for 5 seconds of an immunoblot with the anti-GST antibody.

The lysate was run out on a 12% SDS-polyacrylamide gel with appropriate size markers (see 2.10.1) and controls. The gel was then transferred onto Hybond-P membrane and western blotted using the anti-GST primary antibody (see Table 2.9) and appropriate secondary antibody (see Table 2.10). The 5 second autorad clearly shows each of the interacting proteins at the predicted protein sizes (see Table 8.2). This proved that the pDEST27 interactor constructs were being expressed in the mouse lens epithelial cells and that the cell lysis method (see 2.11.1) was sufficient to release these expressed proteins from the cells so that they could be identified by SDS-PAGE and western blot.

My next step was to use co-immunoprecipitation to investigate whether I could isolate the GST tagged interactor proteins bound to the native PAX6 protein from the mv+ cell line. However, the absence of a preclearing step (see 2.11.3) in my initial experiments meant that there was high cross reactivity with the western blot antibodies meaning that no results could be deciphered. Even with the inclusion of the preclearing step, although the bands corresponding to the purified GST tagged HOMER3 protein could be seen when probed with the anti-GST antibody there was still a smear of signal on the autorads between approximately 20-50kDa. It was discovered that this was because too much anti-GST antibody was being used in the co-immunoprecipitation, and this was causing cross reactivity with the western blot antibodies. By titrating the antibody concentration used in the co-immunoprecipitation of pDEST27_HOMER3_T2 transfected cell lysate I determined that the original 20μg concentration was far too high and that a concentration of 0.1μg was more appropriate (see Figure 8.2).

The autorad in Figure 8.2 clearly shows that the extent of the cross reactivity with the anti-GST antibody light and heavy chains (23kDa and 55kDa respectively) was proportional to the amount of antibody that was used in the co-immunoprecipitation reaction. By using a concentration of 0.1μg of anti-GST antibody in the co-immunoprecipitation reaction, this almost completely eradicated the 55kDa heavy chain cross reactivity band when analysed by western blot. However, even with the antibody titration there was still a protein smear between 20-30kDa, which must be caused by cross reactivity with more than just the co-immunoprecipitation antibody light chain.

I decided that some of the 28kDa protein G (Akerstrom *et al.*, 1987), used to bind the Fc region of the co-immunoprecipitation antibody, was contaminating the samples and causing this protein smear when analysed by western blot. I attempted to minimise the transfer of the protein G sepharose beads by spinning them 3 times after they had been boiled in 2x Laemmli sample

buffer (see 2.11.4). Obviously this method was insufficient to completely remove the protein G from the samples and therefore when analysed by SDS-PAGE and western blot the protein smear persisted whether probed with the anti-GST antibody (see Figure 8.2) or the N terminal PAX6 antibody (see Figure 8.3).

Nevertheless, it is encouraging to see that apart from the cross reactivity with the co-immunoprecipitation antibody subunits and protein G, that the 58.2kDa GST tagged Homer3_T2 fusion protein was purified from non-specific bands (see Figure 8.2) using this co-immunoprecipitation procedure (see 2.11).

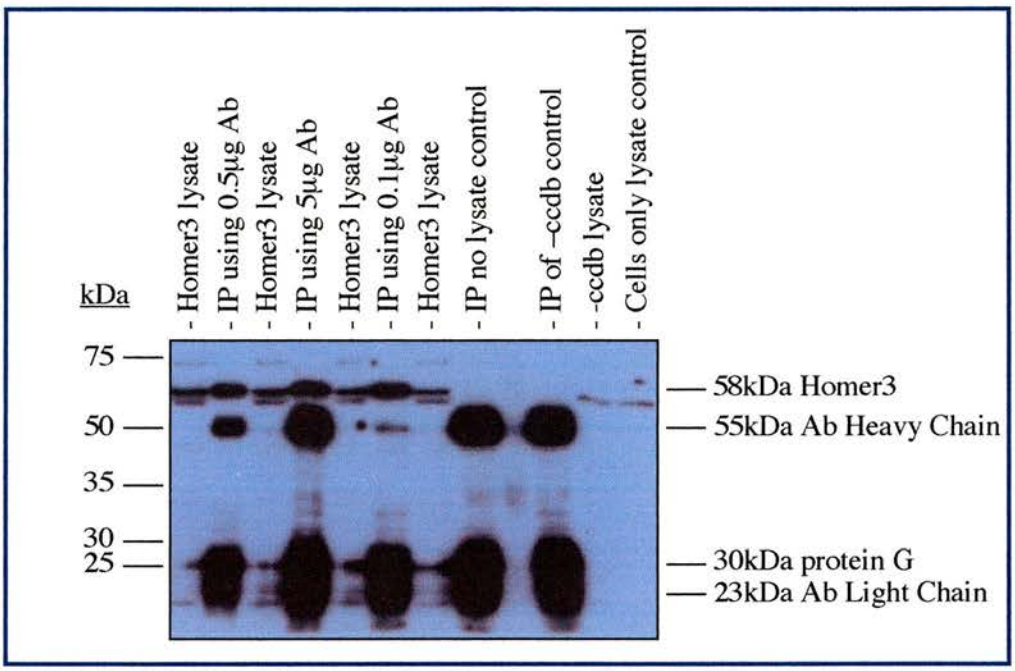



**Figure 8.2 Co-immunoprecipitation of Homer3.**

mv+ cells transfected with pDEST27_Homer3_T2. 200µl of cell lysate was precleared and then co-immunoprecipitated using 0.1µg, 0.5µg and 5µg of anti-GST antibody (see Table 2.9). Lysate and IP samples run on a 12% acrylamide gel. Autorad has been exposed for 10 seconds of an immunoblot with the anti-GST antibody.

Having optimised the amount of antibody that was to be used in future co-immunoprecipitations (0.1µg anti-GST antibody), I transfected mv+ cells with either pDEST27_Homer3_T2, pDEST27_Dncl1_XT31 or pDEST27_Trim11_XA3 and used the resulting cell lysates in co-immunoprecipitation reactions. These samples were then prepared as before to minimise the transfer of protein G beads and then analysed by SDS-PAGE and western blot.

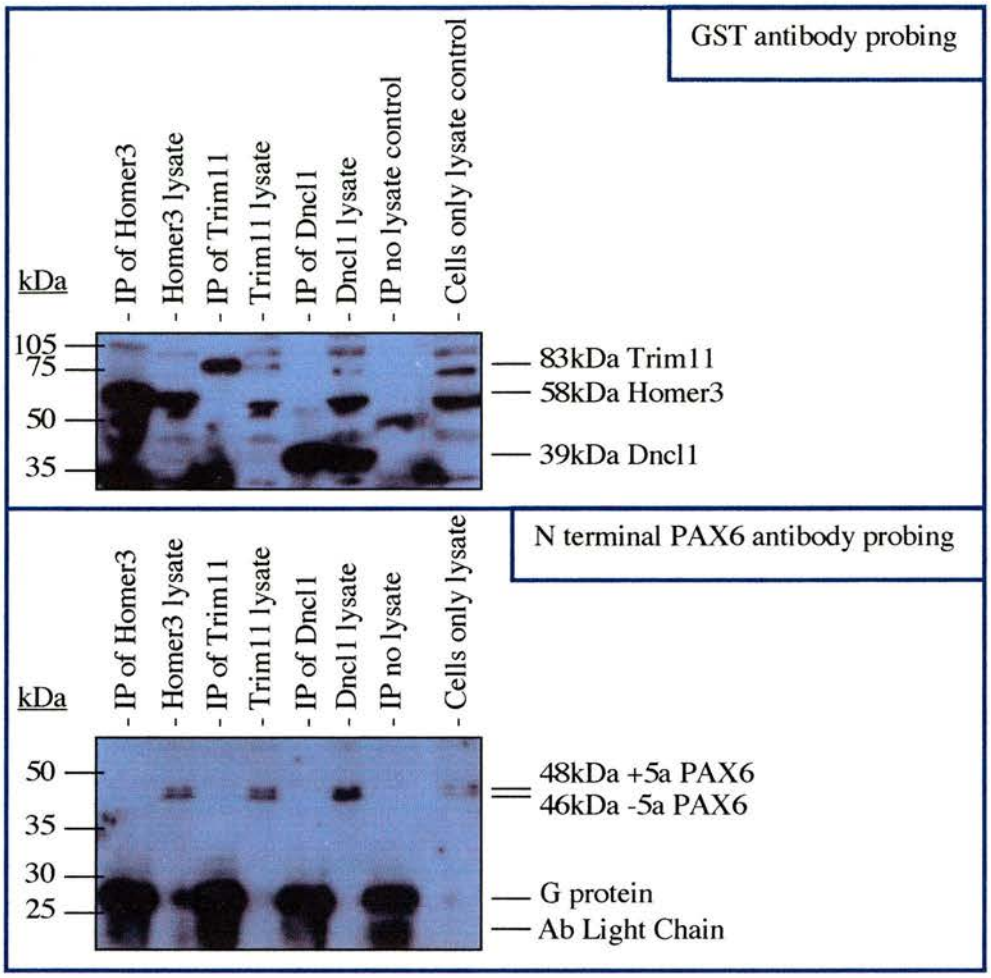



**Figure 8.3 Co-immunoprecipitation of all the interacting proteins.**

mv+ cells transfected with pDEST27_Homer3_T2, pDEST27_Dncl1_XT31 and pDEST27_Trim11_XA3. 200µl of cell lysate was precleared and then co-immunoprecipitated using the anti-GST antibody at a concentrations of 0.1µg (see Table 2.9). Lysate and IP samples were run on 12% acrylamide gels. GST autorad exposed for 4mins and PAX6 autorad exposed for 50mins of an immunoblot with the anti-GST antibody and PAX6 N terminal antibody cocktail respectively. The protein G and Ab light chain smear has been removed from the GST antibody autorad picture.

It can be seen from the top panel of Figure 8.3 that GST tagged interactor proteins could be purified from the crude cell lysate using the anti-GST antibody. There is still some cross reaction with other proteins including protein G and antibody light and heavy chains. Although the protein smear between 20-30kDa was still present when western blotted using the N terminal PAX6 antibody (see Table 2.9) the autorad picture clearly demonstrates that there is no

interaction between the GST tagged interactor proteins and full length PAX6 using this co-immunoprecipitation technique (see Figure 8.3).

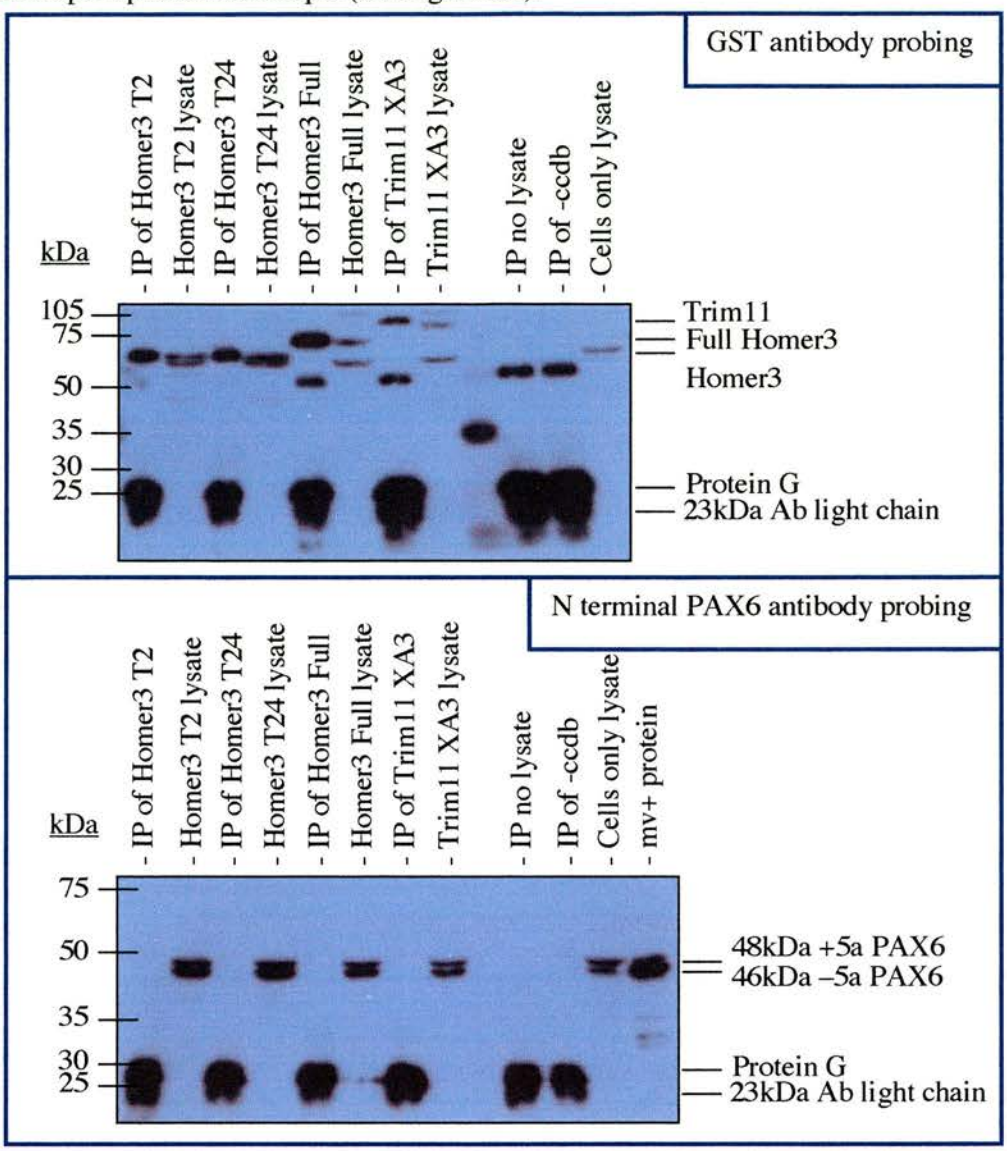



**Figure 8.4 Optimised co-immunoprecipitation of Homer3 and Trim11.**

Co-immunoprecipitation of Homer3_T2, Homer3_T24, Homer3_Full and Trim11_XA3 transfected mv+ cells. 200µl of cell lysate was precleared and then co-immunoprecipitated using 0.1µg of the anti-GST antibody (see Table 2.9). Lysate and IP samples run on 12% acrylamide gels. GST autorad exposed for 1min and PAX6 autorad exposed for 5mins of an immunoblot with the anti-GST antibody and PAX6 N terminal antibody cocktail respectively.

In previous experiments the pDEST27 constructs (see Table 8.1) were transfected into 95% confluent mv+ cells (see 2.9.5); however, after 48hrs incubation at 37°C these cells were often overgrown and in some cases had begun to apoptose (die) and detach from the plate. Therefore, I

decided to transfect the cells at a lower density (50% confluent cells) with the Homer3_T2, Homer3_T24, Homer3_Full and Trim11_XA3 pDEST27 constructs (see Table 8.1). The normal co-immunoprecipitation protocol was then followed and the resulting samples analysed by SDS-PAGE and western blot (see Figure 8.4).

Although there is still some cross reactivity with the antibody heavy chain and the protein smear between 20-30kDa is still evident, it can be clearly seen that all of the Homer3 proteins and the Trim11 protein are purified. However, it can be seen by looking at the western blot with the N terminal PAX6 antibody that there is no interaction with full length PAX6.

## 8.7 Results with C terminal PAX6 antibody.

The co-immunoprecipitation was optimised over the previous set of experiments and using the N terminal PAX6 antibody probing it was identified that using this system none of the GST tagged interacting proteins, when expressed in lens epithelium cells, interacted with native full length PAX6.

The next step of my PhD was to identify whether co-immunoprecipitation could be used to show that these GST tagged interacting proteins would be able to interact with the minus paired domain (-PD) isoform of PAX6, that has been demonstrated to have a cytoplasmic localisation (see Figure 5.10) (Carriere *et al.*, 1995; Zhang *et al.*, 1998). This involved performing the co-immunoprecipitation reactions (see 2.11) on mv+ cells transfected with the GST tagged interacting proteins and the subsequent western blot would be performed using the C terminal PAX6 antibody (see Table 2.9) which is able to detect the –PD PAX6 isoform (see Figure 5.10).

2.5µg of each of the pDEST27 constructs Homer3_T2, Homer3_Full, Dncl1_XT31 and Trim11_A47 (see Table 8.1) were transfected into 50% confluent mv+ cells and the resulting cell lysate was precleared and then co-immunoprecipitated (see 2.11) using 0.1µg of anti-GST antibody. The co-immunoprecipitation samples, transfected cell lysate and relevant controls were then analysed by SDS-PAGE and western blotting using the anti-GST antibody or the C terminal PAX6 antibody (see Table 2.9) and the relevant secondary antibodies (see Table 2.10).

At first the co-immunoprecipitation samples were run out on a 12% SDS polyacrylamide gel which meant that due to the recurrent protein smear between 20-30kDa that the 24kDa –PD PAX6 band, if present in these samples, was obscured (Data not shown). Therefore, this

experiment was repeated and the resulting samples were run out on an 18% SDS polyacrylamide gel (Data not shown). Although this meant that the smaller sized protein bands were better resolved on the gel, after western blot the 24kDa –PD PAX6 band was still obscured in the middle of the protein smear between 20-28kDa, caused by the protein G and antibody light chain.

In an attempt to minimise the transfer of protein G and avoid this protein smear the co-immunoprecipitation reactions were repeated in mini-spin columns (Sigma). When the co-immunoprecipitation reactions were analysed by SDS-PAGE and western blot it was seen that the GST tagged interactor proteins were purified from the crude cell lysate (as in Figure 8.3). However, once again there was severe cross reactivity with the protein G and antibody light chain causing a protein smear between 20-30kDa meaning that I was still unable to asses whether there was any interaction between these GST tagged interactor proteins and the –PD PAX6 isoform (Data not shown).

## 8.8 Results with FLAG tagged –PD PAX6.

I decided to perform this experiment in an attempt to confirm a specific interaction between the 3 interacting proteins (GST tagged, see Table 8.1) and the minus –PD isoform of PAX6 (FLAG tagged, see Table 8.3). However, this experiment is highly contrived as I am over expressing 2 artificially generated fusion proteins in the lens epithelial cell line.

The FLAG tagged –PD PAX6 fusion protein can be expressed in the mouse lens epithelium cell line and can be detected in the cell lysate by western blot using the C terminal PAX6 primary antibody (see Figure 5.10) or the FLAG primary antibody (data not shown, see Tables 2.9 and 2.10) with the correct secondary antibodies. However, the co-immunoprecipitation reactions using this FLAG-PD PAX6 protein were only very preliminary (data not shown) as it was quickly identified that in these reactions the 27kDa FLAG-PD PAX6 fusion protein band would be masked by the smear between 20-30kDa that is caused by the protein G and the antibody light chain (smear illustrated in Figure 8.2).

## 8.9 *In vitro* co-immunoprecipitation.

Due to the problems that were encountered as a result of performing co-immunoprecipitation *in vivo*, I decided to investigate the interaction between the PST domain of PAX6 and the interacting proteins Homer3, Dncl1 and Trim11 using an *in vitro* radioactively labelled protein co-immunoprecipitation technique (see Figure 8.5).

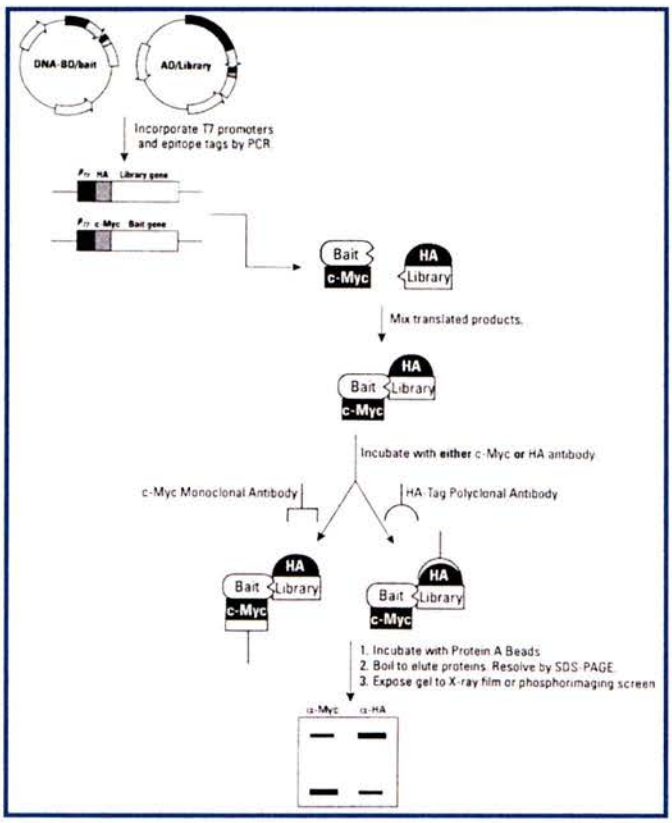



**Figure 8.5 Overview of the BD Matchmaker Co-IP Kit.**

pDBLeu constructs (DNA-BD/bait) are attached to a cMyc epitope tag and pPC86 constructs (AD/Library) to a HA epitope tag (see 2.4.6). These constructs are then translated into $^{35}S$ radioactively labelled protein *in vitro* (see 2.10.7) and the interaction between Homer3, Dncl1 or Trim11 with the PAX6 PST domain is investigated using the Matchmaker Co-IP kit (see 2.11.5).

## 8.10 Cloning Matchmaker co-immunoprecipitation constructs.

In order to be able to use the Matchmaker co-immunoprecipitation system to re-confirm the protein-protein interactions identified in the yeast 2-hybrid system, I needed to modify constructs that represented: the PST domain of PAX6 (pDBLeu_ASNT_norm, see Table 4.1); and each of the interacting proteins Homer3 (pPC86_Homer3_T2), Dncl1 (pPC86_Dncl1_A62) and Trim11 (pPC86_Trim11_A47).

The DNA of these constructs was used in touchdown PCR (see 2.4.6) with the specifically designed primers (see Table 2.5) that will amplify the cDNA inserts as well as introduce an upstream T7 promoter and relevant epitope tag (see Table 8.4). The PCR products generated by the touchdown PCR were checked by running on an agarose gel (2.2.7) and then the remainder

of the PCR was cleaned using the QIAquick PCR purification kit (see 2.2.6). However, when visualising the pDBLeu_ASNT_norm PCR product by agarose gel electrophoresis it was discovered that 2 differently sized PCR products were generated. This was caused by a T7 promoter in the pDBLeu vector, which meant that the forward touchdown PCR primer (ST048) mis-primed causing the additional PCR product. Therefore, the correctly sized PCR product was cut out and purified from the agarose gel (see 2.2.8).

These modified constructs were then translated using the TnT coupled reticulocyte lysate system (see 2.10.7) to generate a radioactively labelled PST domain PAX6 protein with an N terminal cMyc tag and interactor proteins with an N terminal HA tag (see Table 8.4).

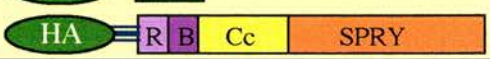
| Modified Construct | Size (aa) | kDa Size | Schematic Diagrams |
|---|---|---|---|
| pDBLeu_ASNT_norm | 168aa | 18.5kDa |   |
| pPC86_Homer3_T2 | 310aa | 33.9kDa |   |
| pPC86_Dncl1_A62 | 148aa | 16.3kDa |   |
| pPC86_Trim11_A47 | 533aa | 58.6kDa |   |

**Table 8.4 Radioactively labelled *in vitro* translated proteins.**

For each of the *in vitro* translated proteins the size in amino acids (aa) and kDa is given for the total protein. The schematic diagrams illustrate the N terminal cMyc or HA antibody tags (orange and green oval respectively) and the domain structure of each protein (see Table 4.1 for PAX6 domains and Table 5.8 for interactor protein domains). The Dncl1 and Trim11 proteins are also illustrated with a linker region (34aa or 41aa respectively) due to the translation of the 5'UTR included in these constructs.

## 8.11 *In vitro* Matchmaker co-immunoprecipitation.

The $^{35}S$ radioactively labelled proteins were then used in the matchmaker co-IP kit (see 2.11.5) along with the cMyc tagged murine p53 and HA tagged SV40 large T antigen positive controls, to investigate *in vitro* protein-protein interactions (see Figure 8.6).

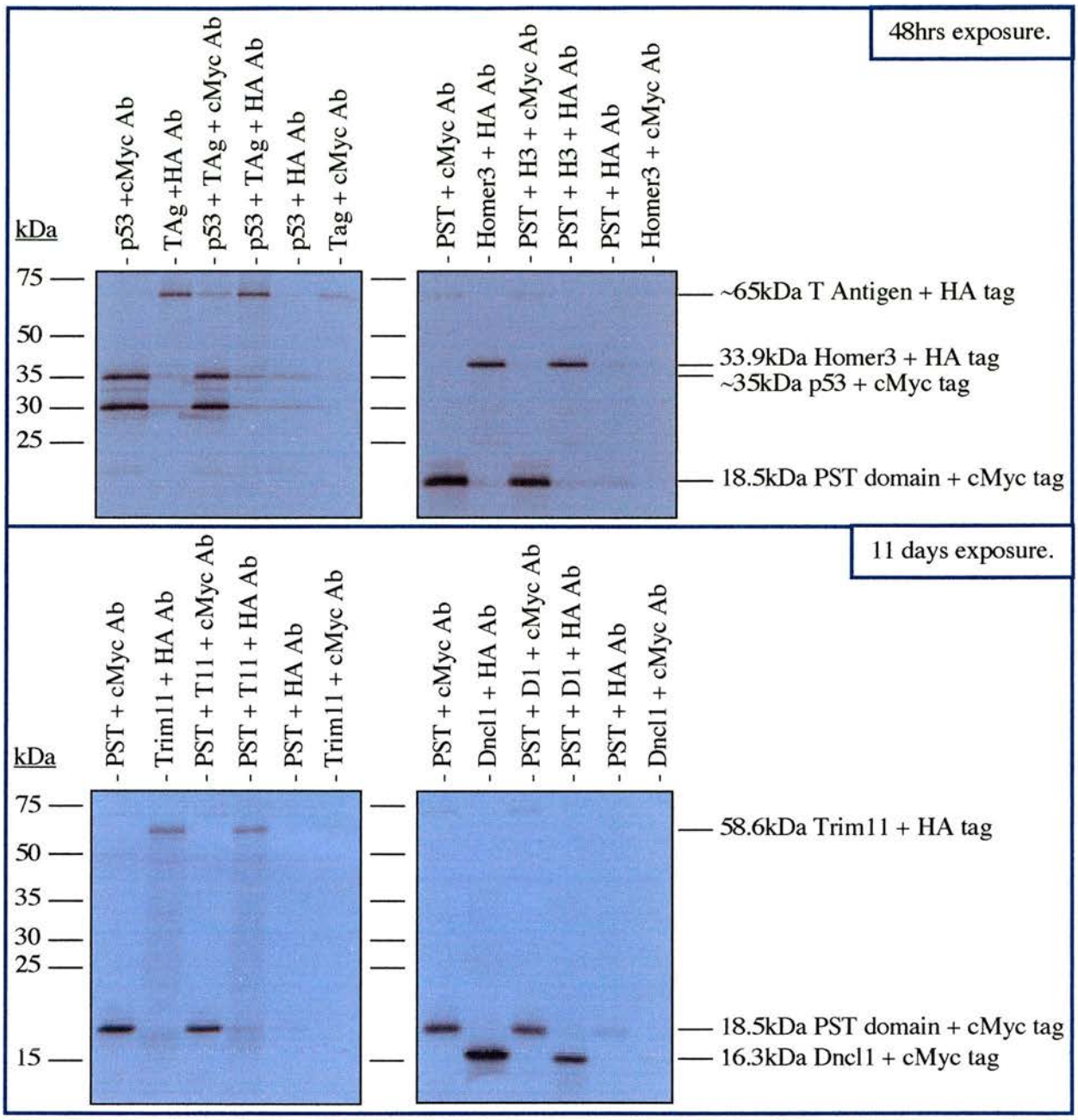



**Figure 8.6 *In vitro* radioactively labelled protein co-immunoprecipitation.**

Each *in vitro* translated radioactively labelled protein was co-immunoprecipitated with both antibodies, cMyc and HA. Possible interacting proteins, including the supplied positive controls p53 and SV40 large T antigen (TAg), were mixed and co-immunoprecipitated with both the cMyc and HA antibodies (see 2.11.5). The resulting samples were then analysed as outlined in 2.11.5 and the resulting radioactive gel exposed to X ray film for either 48hrs or 11days. PST represents the entire PST domain clone, H3, is Homer3, T11 is Trim11 and D1 is Dncl1.

The modified PCR products are successfully translated into the predicted antibody tagged proteins (see Table 8.4) and the cMyc labelled PST domain of PAX6 and the HA labelled interactor proteins can be successfully purified using the corresponding antibody (cMyc and HA respectively) in the co-immunoprecipitation technique (see Figure 8.6).

To check whether there was any cross reactivity, negative controls were performed in which the cMyc tagged PST domain protein was incubated with the HA antibody and the HA tagged interactor proteins were incubated with the cMyc antibody to investigate if these non functional co-immunoprecipitation reactions resulted in the purification of the incubated protein. Each of the HA tagged interacting proteins are not purified with the cMyc antibody. However, it can be seen that the cMyc tagged PAX6 protein (PST) always cross reacts with the HA antibody (see Figure 8.6).

Therefore, when the cMyc tagged PAX6 protein (PST) is mixed with each of the HA tagged interacting proteins and co-immunoprecipitated with the HA antibody it is unsurprising that the result resembles a confirmed interaction. However, due to the cross reactivity of the cMyc tagged PAX6 protein (PST) with the HA antibody it is presumably not a true interaction. This is backed up by the result of incubating the mixed proteins with the cMyc antibody as none of the interactions between the cMyc tagged PAX6 protein (PST) and the HA tagged interacting proteins are re-confirmed (see Figure 8.6).

The strong interaction of the tumour suppressor protein p53 with the SV40 large T antigen (D'Orazi *et al.*, 2002) is used as a positive control in the matchmaker co-immunoprecipitation (see 2.11.5). The cMyc tagged p53 protein can be purified using the cMyc antibody and the HA tagged large T antigen can be purified using the HA antibody. However, when the p53 and T antigen are incubated with the antibody that does not correspond to the proteins epitope tag, it can be seen that there is cross reactivity and weakly purified p53 and T antigen bands are present (see Figure 8.6). This therefore casts doubt as to the validity of the positive interaction seen when the p53 and T antigen proteins are mixed and co-immunoprecipitated with either the cMyc or HA antibody (see Figure 8.6).

This *in vitro* co-immunoprecipitation technique does not re-confirm the interactions between the PST domain of PAX6 and the interacting proteins Homer3, Dncl1 and Trim11 as was hoped. A possible explanation for why could be that these interactions are not strong enough to be identified by this technique, unlike the strong interaction between the p53 protein and the SV40

large T antigen, which is used as a positive control (D'Orazi *et al.*, 2002). Furthermore, the validity of this technique is questionable due to the cross reactivity that occurs in the positive control. Even the confirmation of the known interaction of p53 and the large T antigen is problematic using this technique. However, insufficient time meant that I could not investigate this technique further in an attempt to resolve these problems and confirm possible interactions between PAX6 and Homer3, Dncl1 or Trim11.

## 8.12 Future Studies.

The *in vivo* co-immunoprecipitation has illustrated that the interacting proteins Homer3, Dncl1 and Trim11 do not interact with full length PAX6 using this system. Due to technical difficulties that have already been described in this chapter, it was not possible to determine whether the 3 interacting proteins interact with a minus paired domain isoform of PAX6 *in vivo* or with the PST domain of PAX6 *in vitro*. One way to overcome this problem *in vivo* would be to increase the size of the translated tagged –PD PAX6 fusion protein by adding a larger tag than the FLAG tag (see 8.4). Therefore, if isolated by co-immunoprecipitation the resulting band of the –PD PAX6 fusion protein would be larger than the smear between 20-30kDa caused by the cross reactivity of the protein G and antibody light chain with the western blot antibodies.

### 8.12.1 Matrix assisted laser desorption/ionisation (MALDI) technique.

The MALDI technique utilises a mass spectrometer (MS) to identify unknown proteins by peptide mass fingerprinting. This technique can be used to investigate the identity of the proteins between 20-30kDa that correspond to the protein smear seen when a co-immunoprecipitation reaction is run out on an SDS polyacrylamide gel and western blotted. It will then be possible to determine whether this protein smear contains not only the protein G and antibody light chain contaminants but also the –PD PAX6 isoform isolated due to its interaction with the GST tagged interactor protein in the co-immunoprecipitation reaction.

### 8.12.2 Ciphergen's SELDI ProteinChip System.

Another future experiment involves using the SELDI ProteinChip system (Ciphergen, http://www.ciphergen.com/). As for a normal co-immunoprecipitation the mv+ cell line would be transfected with the relevant GST tagged interactor protein. However, instead of performing the co-immunoprecipitation as described in section 2.11.4 the anti-GST antibody would be

covalently coupled to a pre-activated surface of a ProteinChip array. Then the transfected cell lysate would be applied directly to this ProteinChip array and unbound proteins washed off after a short incubation time to allow interaction of the GST tagged proteins with the coupled antibody. The ProteinChip array would then be analysed using a mass spectrometer and this will hopefully detect not only the antibody and purified GST tagged protein, but also a PAX6 protein isoform. Therefore, this is another possible technique to confirm the interaction of Homer3, Dncl1 and Trim11 with PAX6.

# Chapter 9 Discussion

## 9.1 Discussion.

PAX6 is known to be a developmental transcription factor that is specifically involved in controlling a series of crucial steps in eye development, but has also been shown to be involved in the correct patterning of the brain, including cortical patterning (Bishop *et al.*, 2000), cortical neurogenesis (Heins *et al.*, 2002), specification of neuronal subtypes (Takahashi and Osumi, 2002), and axon guidance (Mastick *et al.*, 1997).

I have shown in Chapter 3 that the final 32 amino acids of the PAX6 PST domain are distinct from the rest of the PST domain in having a highly conserved amino acid sequence and a specific secondary structure. I have therefore, classified this region as a separate domain, the C terminal domain. At the time of undertaking this work I was unaware that similar characterisation had already been performed (Mikkola *et al.*, 1999). However, it is reassuring that the same region has already been described and the same secondary structure predicted although this is the first time that this region has been specifically described as a separate domain. The question that needed to be addressed in light of the description of this evolutionary conserved C terminal domain was "Does the C terminal domain of PAX6 have a specific function?"

This question led to the discovery that the PST domain of PAX6 interacts with three proteins Homer3, Dncl1 and Trim11, as determined by yeast 2-hybrid screening of a mouse brain cDNA library, and this is the first report of an interaction between the PST domain of PAX6 and other proteins. More intriguing is the fact that two of these interactions (with Homer3 and Dncl1) are specifically dependent on the newly characterised C terminal domain of PAX6.

This has been the first systematic screen for proteins that interact with PAX6. Although protein interactions with the PAX family of transcription factors have been found, including several interactions between Pax and Sox proteins, these have not been uncovered by systematic library screens. As there seems to be a possible functional link between the PAX and SOX protein families (Chi and Epstein, 2002), then it is possible that there could be a link between PAX proteins and either the HOMER or TRIM protein families or subunits of the dynein motor complex. Therefore, these interactions may influence the regulation and distribution of these transcription factors in development and mutation may result in human disease caused by an inability of these proteins to carry out their normal functions.

It is interesting to hypothesise about a possible reason for the interaction between Homer3, Dncl1, Trim11 and PAX6. In the sections that follow (9.1.1 to 9.3) for the sake of simplicity I will use generic lower case gene and protein names, even when I am referring to the human gene or protein.

### 9.1.1 Trim11.

The function of Trim11 is still unclear although it is possible that it may regulate the expression of Pax6 either by specific ubiquitin-mediated degradation of the protein due to Trim11 association (Niikura *et al.*, 2003) or by correctly positioning Pax6 in a macromolecular complex (Borden, 2000) so that it functions normally. As the PST domain of Pax6 is involved in the transactivational role of Pax6 then it is possible that Trim11, due to its interaction around the SYSC motif in Pax6 (green box, see Figure 3.1), allows the correct positioning of the components of the transcriptional machinery, allowing Pax6 to execute its normal function. However, due to the prior link of the Trim11 proteins with ubiquitination (Niikura *et al.*, 2003) then it is more likely that the interaction of Trim11 with Pax6 could potentially interfere with its transcriptional activity and couple the Pax6 protein to ubiquitin degradation pathways. In this way the levels of Pax6 could be regulated by Trim11 in the brain.

### 9.1.2 Dncl1.

Dncl1 has been shown to be a component of microtubule (King *et al.*, 1996) and actin (Espindola *et al.*, 2000) based molecular motors and is thought to be involved in the coupling of proteins to these motors and their subsequent movement around the cell. This is very important when considering the interaction between Dncl1 and Pax6 as it suggests a mechanism of how Pax6 is translocated throughout the cell. Dncl1 has also been shown to bind to nNOS (Rodriguez-Crespo *et al.*, 1998) and GKAP (Rodriguez-Crespo *et al.*, 2001) and it is thought that Dncl1, due to its association with cytoplasmic dynein, is involved in axonal transport of these proteins. Interestingly, the post-synaptic density has been shown to include both these proteins coupled to NMDA receptors through the PSD-95 scaffolding protein and so it is likely that both nNOS and GKAP are transported between the PSD and the nucleus via the dynein molecule, due to their interaction with Dncl1. This is fascinating as it is possible that Dncl1 specifically functions to transport Pax6 from a PSD location to the nucleus in developing neurons.

Lissencephaly causes severe brain defects due to problems in neuronal migration that produce brains lacking the normal convoluted surface (Vallee *et al.*, 2001). Although the most common phenotype of these syndromes is the severe brain defect that results in mental retardation and normally early death caused by uncontrollable seizures (Vallee *et al.*, 2001), there are also related eye defects that occur in type II lissencephaly. However, at present it is only lissencephaly type I that is thought to result from mutations in the *LIS1* gene, which cause haploinsufficiency (Vallee *et al.*, 2001). The LIS1 protein is known to interact with the cytoplasmic dynein motor protein and it is thought that the brain defects seen in lissencephaly might be caused by a malfunctioning of dynein regulation (Vallee *et al.*, 2001). Therefore, it is interesting to hypothesise that other dynein-interacting proteins may also be involved in brain patterning and neuronal migration, and that mutant forms of these interactors (such as the Pax6 mutants described in 1.7.6) may cause specific brain defects.

*Dncl1* mutations in *Drosophila* are lethal but are characterised by severe neuronal defects, which has led to the suggestion that Dncl1 may be involved in axon formation (Jacob *et al.*, 2000). It is therefore interesting to consider that the interaction between Dncl1 and Pax6 may facilitate the transport of Pax6 from the PSD to the nucleus in order to activate gene expression (Herzig *et al.*, 2000). Therefore, due to the mutation of Dncl1, the function of Pax6 in neuronal development will be undirected, possibly contributing to these severe neuronal defects of axon formation in *Drosophila* Dncl1 mutants. However, it must be remembered that it is not surprising that the *Drosophila* Dncl1 mutants cause severe neuronal defects as Dncl1 has been implicated in the transport of many post-synaptic proteins which link post-synaptic receptors to intracellular signaling cascades.

Although the characterisation of 2 potential Dncl1 binding regions in Pax6 was based on very weak regions of homology with other identified Dncl1 binding sites (see 5.6.2) it was intriguing to consider this when analysing the results of the yeast 2-hybrid pairwise results. It can be seen that the entire PST domain interacts with Dncl1; however, the SYSC truncated peptide does not interact with Dncl1 (see Table 6.2). Although the interaction between Dncl1 and the C terminal domain could not be replicated, when the C terminal domain was removed from the entire PST domain construct then there was again no interaction with Dncl1. I have already explained how I think the conformation of the entire PST domain is important in carrying out its transactivation function (see 7.3) and in interacting with other proteins. It is possible that this correct conformation means that these 2 Dncl1 binding regions are correctly positioned so that Dncl1 and Pax6 can interact.

It is intriguing to consider that an interaction also occurs between Dncl1 and the PST domain of other PAX proteins, and could potentially provide a mechanism for their cellular distribution. Therefore, mutations in the PST domain of PAX proteins that alter the Dncl1 binding sites would cause these proteins to be incorrectly distributed within the cell, which may provide an explanation for developmental defects or human disease.

### 9.1.3 Homer3.

Homer3 is located within the postsynaptic density (PSD) (Dick *et al.*, 1996; Xiao *et al.*, 1998) and regulates the metabotropic glutamate receptors mGluR1 and mGluR5 (Sheng and Kim, 2002) by binding to a proline rich motif in the C terminal tail of the receptors (Tu *et al.*, 1998). It has also been shown that Homer3 is central in producing a scaffold that can potentially link these type-I mGluRs to the endoplasmic reticulum through IP3Rs and RyRs and has also been shown to be involved in linking mGluRs with NMDA receptors due to a large complex involving Homer3, Shank, PSD-95 and GKAP. To date there has only been one reported interaction with the coiled coil domain of Homer3, other than with other Homer family members to facilitate multimerisation, and this was with S8 ATPase, a protein proposed to be involved in Homer3 mediated degradation of mGluRs (Rezvani *et al.*, 2003). When dissociated from the Homer3 scaffolding complex it was found that S8 ATPase translocated to the nucleus (Rezvani *et al.*, 2003). Therefore, it is interesting to speculate that Pax6 could be sequestered in the PSD due to its interaction with Homer3 and then transported to the nucleus via its interaction with Dncl1.

Activation of NMDA receptors stimulates mitogen-activated protein kinase (MAPK), which is known to phoshphorylate Pax6 (Mikkola *et al.*, 1999). Therefore, it is interesting to envisage a situation where synaptic activation leads to the stimulation of the NMDA receptor which in turn activates MAPK. As Pax6 is sequestered in the PSD due to its interaction with Homer3, then it is feasible that the activated MAPK could phosphorylate Pax6, producing a modified protein that has an increased transcription activity. As it is probable that synaptic activation would also cause the upregulation of Homer1a, a Homer protein that antagonises the scaffolding function of Homer3, then the activated Pax6 could be released from the PSD and be transported to the nucleus, due to its interaction with Dncl1. The activated Pax6 could then be transported to sites of active gene expression in the nucleus to allow transcription of Pax6 target genes. In this way it is possible that the resulting transcriptional activation may be involved in the development of proper synaptic connections, spine morphogenesis or neuronal migration in the developing brain. In the adult brain it is possible that this same signalling cascade could be involved in maintaining

normal synaptic function or in the production of long term potentiation, which is thought to be involved in learning and memory.

Further evidence to back up this hypothesis comes from the investigation of *Pax6* mutant rats, which led the authors to suggest that Pax6 functions to regulate the cytoskeletal organisation during polarisation and migration of CNS neurons (Yamasaki *et al.*, 2001). In wild-type rats Pax6 mRNA and protein are both expressed throughout differentiation and migration of cerebellar granule cells in these cells (Yamasaki *et al.*, 2001). Although Pax6 is not expressed in mutant cells it was shown that the gene cascade required for granule cell differentiation was still produced (Yamasaki *et al.*, 2001). It was also shown that it was specifically the inability of Pax6 to carry out its gene expression function in the nucleus that causes problems in neuron migration and cellular morphogenesis due to abnormal polarisation of the cerebellar granule cells. Therefore, it was hypothesised that Pax6 acts as an intrinsic upstream factor that stimulates downstream target genes involved in the elongation of bipolar axons and their orderly migration in the developing brain (Yamasaki *et al.*, 2001).

Specifically, it has been suggested that Pax6 could potentially be involved in the control of the cytoskeletal networks during the polarisation of neural cells in the CNS and the role of Pax6 might be to activate gene expression of target genes that are involved in the organisation of cytoskeletal proteins in granule cells (Yamasaki *et al.*, 2001). It has been suggested that Pax6 might potentially be involved in sustaining the formation of bipolar processes in tangential migration of the cerebellar granule cells, which are guided, by the leading processes that contact the axon fibres that have already formed (Yamasaki *et al.*, 2001).

Therefore, it is possible that although these mutant rats seem to have the correct gene cascade required for granule cell differentiation but are lacking Pax6, then perhaps the inability to activate Pax6 dependent gene expression due to stimulation by extracellular cues in the developing neurons, leads to the reported inability to reorganise the cytoskeleton and the enlarged growth cones (Yamasaki *et al.*, 2001). This report represented the first description of how Pax6 functions in regulating neuronal morphogenesis and migration by altering the cytoskeletal network (Yamasaki *et al.*, 2001); however, I have now provided an additional link to this hypothesis that suggests that the activation of this Pax6 mediated genetic cascade is caused by extracellular cues, relayed through cell surface receptors.

This function of Pax6 in the regulation of neurite outgrowth and cell migration has also been proposed to be further regulated due to the inhibitory effects caused by the binding of a protein called mDia to Pax6 (Tominaga *et al.*, 2002). Therefore, this shows how the stimulation of neuronal cells can cause the activation of gene cascades that allow proper neuronal migration and alteration of cellular morphogenesis and how this has to be negatively regulated so that the cytoskeleton can be re-organised.

Due to this possible role of Pax6 in regulating the cytoskeletal network it is therefore interesting to consider the lethal mouse mutant, *dilute*, which is characterised by a convulsive neurological disorder (Petralia *et al.*, 2001). These mice are lacking the myosin-Va transport protein, involved in the organisation of the cytoskeleton, which has been tentatively suggested to be a potential downstream target of Pax6 mediated gene activation (Petralia *et al.*, 2001). In these mice it is thought that the convulsive disorder is due to the lack of recruitment of IP3R laden endoplasmic reticulum into the dendritic spines even though the distribution of mGluRs, Homer and Dncl1 is not affected (Petralia *et al.*, 2001). This suggests that there is no calcium release from these dendritic spines which is thought to affect the long term depression of these synapses leading to the convulsive disorders (Petralia *et al.*, 2001).

A specific aniridia patient that has the C terminal Pax6 mutation X423L, was also found to have polymicrogyria, a rare brain defect thought to be caused by disruption of neuronal migration (Mitchell *et al.*, 2003), and suffered from epilepsy. It is interesting to assume that in this patient the Pax6 mutation is sufficient to dysregulate the normal Pax6 mediated signalling cascade, due to the inability of the mutant protein to interact with Homer3 and Dncl1 (see Table 6.4). This would mean that due to the failure of Pax6 to activate gene expression that the production of cytoskeletal organisation proteins such as myosin-Va could be reduced or abolished. Therefore, the lack of myosin-Va could potentially explain the phenotype of epilepsy and seizures that is found in this patient. Alternatively, the epileptic phenotype could be an associated phenotype that is specifically caused by the improper neuronal patterning of the brain that results in the polymicrogyria in this patient.

Although the research by Yamasaki *et al.*, (2001) was focused on cerebellar granule cells it was suggested that the regulation of migration by Pax6 may also involve other neuronal cells that utilise leading processes in migration in the cerebellar system (Yamasaki *et al.*, 2001). This fits nicely with similar findings in *Pax6* null mice, where late born cortical neurons show errors in migration and accumulate to create an expanded ventricular zone (Mastick *et al.*, 1997).

Therefore, this inability of neurons that lack Pax6 to migrate could explain why there is no innervation between the thalamus and the cortex in *Pax6* null mice.

A similar problem in neuronal migration and incorrect neuronal connections due to the inability of Pax6 to interact with Homer3 and/or Dncl1 could potentially explain the brain associated phenotypes of aniridia patients with C terminal domain Pax6 mutations such as 1615del10 (WHEDA) or the X423L mutation. The 1615del10 mutation is associated with a neurobehavioural phenotype (Heyman *et al.*, 1999) and the X423L mutation has been associated with a defect in olfactory function in one patient (Sisodiya *et al.*, 2001) and polymicrogyria in 2 patients (Mitchell *et al.*, 2003). Therefore, it is possible that all of these brain anomalies could be caused by incorrect neuronal patterning or connections caused by the defect in the Pax6 mediated gene activation involved in neuronal migration and morphogenesis.

Although the above discussion focuses on the role of Pax6 and the isolated interactors in the developing brain, there is still a connection to the eye. It has been shown that specific expression of Pax6 is required for the correct direction of axons from the retina to the superior colliculus, which mediates ocular functions (Baumer *et al.*, 2002). Therefore, the combination of Pax6, Homer3 and Dncl1 may influence the connection of the neuroretina to discrete regions of the visual centres of the brain due to the combined function in regulating neuronal migration. In addition, expression of Pax6 in the developing neuroretina is important for normal patterning within the developing eye (Baumer *et al.*, 2002). If the C terminal domain is mutated and the normal interaction with Homer3 and Dncl1 is disrupted then normal patterning of the eye and the visual centres of the brain may also be altered.

## 9.2 Interesting future studies.

Using the yeast 2-hybrid pairwise screens I showed that full length Pax6 (-5a Pax6 isoform) was unable to interact with the three interacting proteins (see Table 6.2) or transactivate reporter gene expression (data not shown). This could potentially be due to a conformational change caused by the attachment of the large N terminal GAL4 DNA binding domain. Alternatively, the inclusion of the DNA binding domains of Pax6 could themselves introduce a direct negative effect, as already shown with inclusion of the homeodomain of Pax6 in transactivation potential experiments resulting in the inhibition of the transactivation ability of fusion proteins (Tang *et al.*, 1998; Mikkola *et al.*, 1999). It was proposed that this direct negative effect of the DNA binding domains could be caused by interfering with the transactivation ability of the PST

domain or by directing the protein to bind to other yeast chromosomal regions other than the GAL4 promoter site (Mikkola *et al.*, 1999).

However, it was also determined that Homer3, Dncl1 and Trim11 do not interact with full length Pax6 isoforms in lens epithelial cells, by means of *in vivo* co-immunoprecipitation. Therefore, it is possible that these three proteins interact with the predominantly cytoplasmic –PD Pax6 isoform and it will be interesting to determine whether an interaction between this Pax6 isoform and the 3 interactors can be confirmed using the yeast 2-hybrid system, although similar technical problems as experienced with full length Pax6 may occur.

Following on from this, obviously the highest priority for future studies will be to positively confirm an interaction between Pax6 and these three interactors Homer3, Dncl1 and Trim11 *in vivo*. I have already mentioned in Chapter 8 that one possible addition to the experiments that I have already performed would be to increase the overall size of the –PD Pax6 fusion protein so that co-transfection studies will avoid the problem of cross reactivity of the protein G and antibody light chain with the western blot antibodies.

It is appealing to think that the –PD Pax6 isoform might have a specific function in neuronal cells due to a possible interaction with Homer3, Dncl1 or Trim11. It has already been shown in *Caenorhabditis elegans* that only the ortholog of the –PD Pax6 isoform is expressed in ray cells of the male tail and neurons in the preanal ganglion and during development this isoform translocates from the cytoplasm to the nucleus possibly due to lineage specific or spatial cues during differentiation of these cells (Zhang *et al.*, 1998). Therefore, it is possible that if the –PD Pax6 isoform does interact with Homer3 and Dncl1 then it may be that this specific isoform is sequestered in the post-synaptic density until dissociation of the scaffolding network allows transport to the nucleus via Dncl1. A novel function of the –PD Pax6 isoform may be to cause changes in gene expression related to synaptic development and activation, possibly indicating a role in learning and memory.

I think it would also be worthwhile to investigate the *in vivo* co-immunoprecipitation of Pax6 and the three interacting proteins in various cell types including specific cell lines derived from neural tissues such as the retina and the brain. The subcellular localisation of Pax6 isoforms has already been proposed to be variable between cell types and so these experiments could be used to specifically ascertain whether this hypothesis is true and if this affects the results of *in vivo* co-immunoprecipitation.

I think it is evident from the transactivation studies that a specific conformation is required for Pax6 to produce its transactivation ability and that the C terminal domain is pivotal in producing this structure. It is interesting to think that this interaction is not only important for the transactivation ability of Pax6 but also that this same conformation could potentially be required for the interaction of the PST domain with the three interacting proteins (see 7.3) and that both transactivation and interaction with these interactors is altered by C terminal domain Pax6 mutations. Therefore, on confirmation of the interaction between these proteins and Pax6, another potentially interesting follow up experiment would be to further characterise the region within Pax6 that is responsible for its interaction with these proteins. This could be carried out by utilising deletion analysis and site directed mutagenesis to determine the important amino acids within the C terminal domain of Pax6 for the formation of the interaction complex.

Having investigated the literature on Homer3, I have found that it shows a punctate expression pattern along developing neurons (Xiao *et al.*, 1998; Diagana *et al.*, 2002; Shiraishi *et al.*, 2003). It would therefore be interesting to perform fluorescent localisation studies in neuronal cells to see if this characteristic punctate pattern of Homer3 is reproduced by Pax6. It would also be interesting to confirm that Pax6 does indeed localise to the post synaptic density in neurons and to further investigate which isoform of Pax6 localises to this region with Homer3.

It would also be interesting to see whether it is possible to observe the co-expression of Dncl1 and Pax6 in developing neurons and to determine whether stimulation of neurons causes a translocation of Pax6 from the PSD to the cell nucleus to confirm my hypothesis of the functional significance of this interaction.

A possible interaction between Dncl1 and Homer3 is implied by the observation of a punctate pattern of expression of Dncl1 in dendritic shafts (Naisbitt *et al.*, 2000), which is similar to that of Homer3 (Xiao *et al.*, 1998; Diagana *et al.*, 2002; Shiraishi *et al.*, 2003). Furthermore, Dncl1 has been reported to interact with other PDZ containing proteins such as GKAP and nNOS that are localised within the PSD. Therefore, it would be interesting to see if there was an interaction between Dncl1 and Homer3 and determine whether Homer3 is transported between the PSD and the nucleus due to its interaction with Dncl1.

## 9.3 Concluding remarks.

In conclusion, although speculative at present, it is exciting to think that the interaction of these three proteins with the PST domain of Pax6 represents the first functionally significant protein interactions with this region of Pax6. The importance of the C terminal domain of Pax6 in some of these interactions justifies its description as a separate domain, which is involved in producing a specific conformation. The function of Pax6 in combination with the isolated interactors has been hypothesised to provide a model of how a neuron can change its pattern of gene expression in response to receptor stimuli. Through these interactors Pax6 may regulate aspects of neuronal development such as axon guidance and could also have a role in learning and memory.

# Chapter 10 Bibliography

Aiba, A., Kano, M., Chen, C., Stanton, M. E., Fox, G. D., Herrup, K., Zwingman, T. A. and Tonegawa, S. (1994). Deficient cerebellar long-term depression and impaired motor learning in mGluR1 mutant mice. *Cell* **79** (2), 377-88.

Akerstrom, B., Nielsen, E. and Bjorck, L. (1987). Definition of IgG- and albumin-binding regions of streptococcal protein G. *J Biol Chem* **262** (28), 13388-91.

Alonso, C., Miskin, J., Hernaez, B., Fernandez-Zapatero, P., Soto, L., Canto, C., Rodriguez-Crespo, I., Dixon, L. and Escribano, J. M. (2001). African swine fever virus protein p54 interacts with the microtubular motor complex through direct binding to light-chain dynein. *J Virol* **75** (20), 9819-27.

Ango, F., Prezeau, L., Muller, T., Tu, J. C., Xiao, B., Worley, P. F., Pin, J. P., Bockaert, J. and Fagni, L. (2001). Agonist-independent activation of metabotropic glutamate receptors by the intracellular protein Homer. *Nature* **411** (6840), 962-5.

Aota, S., Nakajima, N., Sakamoto, R., Watanabe, S., Ibaraki, N. and Okazaki, K. (2003). Pax6 autoregulation mediated by direct interaction of Pax6 protein with the head surface ectoderm-specific enhancer of the mouse Pax6 gene. *Dev Biol* **257** (1), 1-13.

Ariniello, L. (2002). Brain Development. *In Brain Facts: A primer on the brain and nervous system.* J. Carey. (The Society for Neuroscience), 8-11.

Axton, R., Hanson, I., Danes, S., Sellar, G., van Heyningen, V. and Prosser, J. (1997). The incidence of PAX6 mutation in patients with simple aniridia: an evaluation of mutation detection in 12 cases. *J Med Genet* **34** (4), 279-86.

Azuma, N., Nishina, S., Yanagisawa, H., Okuyama, T. and Yamada, M. (1996). PAX6 missense mutation in isolated foveal hypoplasia. *Nat Genet* **13** (2), 141-2.

Azuma, N. and Yamada, M. (1998). Missense mutation at the C terminus of the PAX6 gene in ocular anterior segment anomalies. *Invest Ophthalmol Vis Sci* **39** (5), 828-30.

Azuma, N., Yamaguchi, Y., Handa, H., Hayakawa, M., Kanai, A. and Yamada, M. (1999). Missense mutation in the alternative splice region of the PAX6 gene in eye anomalies. *Am J Hum Genet* **65** (3), 656-63.

Azuma, N., Yamaguchi, Y., Handa, H., Tadokoro, K., Asaka, A., Kawase, E. and Yamada, M. (2003). Mutations of the PAX6 gene detected in patients with a variety of optic-nerve malformations. *Am J Hum Genet* **72** (6), 1565-70.

Ball, L. J., Jarchau, T., Oschkinat, H. and Walter, U. (2002). EVH1 domains: structure, function and interactions. *FEBS Lett* **513** (1), 45-52.

Balling, R., Deutsch, U. and Gruss, P. (1988). undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1. *Cell* **55** (3), 531-5.

Baum, L., Pang, C. P., Fan, D. S., Poon, P. M., Leung, Y. F., Chua, J. K. and Lam, D. S. (1999). Run-on mutation and three novel nonsense mutations identified in the PAX6 gene in patients with aniridia. *Hum Mutat* **14** (3), 272-3.

Baumer, N., Marquardt, T., Stoykova, A., Ashery-Padan, R., Chowdhury, K. and Gruss, P. (2002). Pax6 is required for establishing naso temporal and dorsal characteristics of the optic vesicle. *Development* **129** (19), 4535-45.

Beneken, J., Tu, J. C., Xiao, B., Nuriya, M., Yuan, J. P., Worley, P. F. and Leahy, D. J. (2000). Structure of the Homer EVH1 domain-peptide complex reveals a new twist in polyproline recognition. *Neuron* **26** (1), 143-54.

Bernier, G., Vukovich, W., Neidhardt, L., Herrmann, B. G. and Gruss, P. (2001). Isolation and characterization of a downstream target of Pax6 in the mammalian retinal primordium. *Development* **128** (20), 3987-94.

Bishop, K. M., Goudreau, G. and O'Leary, D. D. (2000). Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. *Science* **288** (5464), 344-9.

Bishop, K. M., Rubenstein, J. L. and O'Leary, D. D. (2002). Distinct actions of Emx1, Emx2, and Pax6 in regulating the specification of areas in the developing neocortex. *J Neurosci* **22** (17), 7627-38.

Bonifer, C. (1999). Long-distance chromatin mechanisms controlling tissue-specific gene locus activation. *Gene* **238** (2), 277-89.

Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Solomon, E. and Freemont, P. S. (1996). In vivo and in vitro characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia protooncoprotein PML. *Proc Natl Acad Sci U S A* **93** (4), 1601-6.

Borden, K. L. (1998). RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem Cell Biol* **76** (2-3), 351-8.

Borden, K. L. (2000). RING domains: master builders of molecular scaffolds? *J Mol Biol* **295** (5), 1103-12.

Brakeman, P. R., Lanahan, A. A., O'Brien, R., Roche, K., Barnes, C. A., Huganir, R. L. and Worley, P. F. (1997). Homer: a protein that selectively binds metabotropic glutamate receptors. *Nature* **386** (6622), 284-8.

Busslinger, M., Klix, N., Pfeffer, P., Graninger, P. G. and Kozmik, Z. (1996). Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc Natl Acad Sci U S A* **93** (12), 6129-34.

Byers, P. H. (2002). Killing the messenger: new insights into nonsense-mediated mRNA decay. *J Clin Invest* **109** (1), 3-6.

Caric, D., Gooday, D., Hill, R. E., McConnell, S. K. and Price, D. J. (1997). Determination of the migratory capacity of embryonic cortical cells lacking the transcription factor Pax-6. *Development* **124** (24), 5087-96.

Carriere, C., Plaza, S., Martin, P., Quatannens, B., Bailly, M., Stehelin, D. and Saule, S. (1993). Characterization of quail Pax-6 (Pax-QNR) proteins expressed in the neuroretina. *Mol Cell Biol* **13** (12), 7257-66.

Carriere, C., Plaza, S., Caboche, J., Dozier, C., Bailly, M., Martin, P. and Saule, S. (1995). Nuclear localization signals, DNA binding, and transactivation properties of quail Pax-6 (Pax-QNR) isoforms. *Cell Growth Differ* **6** (12), 1531-40.

Chao, L. Y., Mishra, R., Strong, L. C. and Saunders, G. F. (2003). Missense mutations in the DNA-binding region and termination codon in PAX6. *Hum Mutat* **21** (2), 138-45.

Chi, N. and Epstein, J. A. (2002). Getting your Pax straight: Pax proteins in development and disease. *Trends Genet* **18** (1), 41-7.

Chisholm, A. D. and Horvitz, H. R. (1995). Patterning of the Caenorhabditis elegans head region by the Pax-6 family member vab-3. *Nature* **377** (6544), 52-5.

Chow, R. L., Altmann, C. R., Lang, R. A. and Hemmati-Brivanlou, A. (1999). Pax6 induces ectopic eyes in a vertebrate. *Development* **126** (19), 4213-22.

Crepieux, P., Kwon, H., Leclerc, N., Spencer, W., Richard, S., Lin, R. and Hiscott, J. (1997). I kappaB alpha physically interacts with a cytoskeleton-associated protein through its signal response domain. *Mol Cell Biol* **17** (12), 7375-85.

Cvekl, A. and Piatigorsky, J. (1996). Lens development and crystallin gene expression: many roles for Pax-6. *Bioessays* **18** (8), 621-30.

Czerny, T. and Busslinger, M. (1995). DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Mol Cell Biol* **15** (5), 2858-71.

D'Orazi, G., Cecchinelli, B., Bruno, T., Manni, I., Higashimoto, Y., Saito, S., Gostissa, M., Coen, S., Marchetti, A., Del Sal, G., Piaggio, G., Fanciulli, M., Appella, E. and Soddu, S. (2002). Homeodomain-interacting protein kinase 2 phosphorylates p53 at Ser 46 and mediates apoptosis. *Nat Cell Biol* **4** (1), 11-9.

De Bartolomeis, A. and Iasevoli, F. (2003). The homer family and the signal transduction system at glutamatergic postsynaptic density: potential role in behavior and pharmacotherapy. *Psychopharmacol Bull* **37** (3), 51-83.

Devriendt, K., Matthijs, G., Van Damme, B., Van Caesbroeck, D., Eccles, M., Vanrenterghem, Y., Fryns, J. P. and Leys, A. (1998). Missense mutation and hexanucleotide duplication in the PAX2 gene in two unrelated families with renal-coloboma syndrome (MIM 120330). *Hum Genet* **103** (2), 149-53.

Diagana, T. T., Thomas, U., Prokopenko, S. N., Xiao, B., Worley, P. F. and Thomas, J. B. (2002). Mutation of Drosophila homer disrupts control of locomotor activity and behavioral plasticity. *J Neurosci* **22** (2), 428-36.

Dick, T., Ray, K., Salz, H. K. and Chia, W. (1996). Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in Drosophila melanogaster. *Mol Cell Biol* **16** (5), 1966-77.

Dimanlig, P. V., Faber, S. C., Auerbach, W., Makarenkova, H. P. and Lang, R. A. (2001). The upstream ectoderm enhancer in Pax6 has an important role in lens induction. *Development* **128** (22), 4415-24.

Duncan, M. K., Haynes, J. I., 2nd, Cvekl, A. and Piatigorsky, J. (1998). Dual roles for Pax-6: a transcriptional repressor of lens fiber cell-specific beta-crystallin genes. *Mol Cell Biol* **18** (9), 5579-86.

Duncan, M. K., Kozmik, Z., Cveklova, K., Piatigorsky, J. and Cvekl, A. (2000). Overexpression of PAX6(5a) in lens fiber cells results in cataract and upregulation of (alpha)5(beta)1 integrin expression. *J Cell Sci* **113 ( Pt 18)**, 3173-85.

Ehlers, M. D. (2002). Molecular morphogens for dendritic spines. *Trends Neurosci* **25** (2), 64-7.

Engelkamp, D., Rashbass, P., Seawright, A. and van Heyningen, V. (1999). Role of Pax6 in development of the cerebellar system. *Development* **126** (16), 3585-96.

Epstein, E., Sela-Brown, A., Ringel, I., Kilav, R., King, S. M., Benashski, S. E., Yisraeli, J. K., Silver, J. and Naveh-Many, T. (2000). Dynein light chain binding to a 3'-untranslated sequence mediates parathyroid hormone mRNA association with microtubules. *J Clin Invest* **105** (4), 505-12.

Epstein, J. A., Glaser, T., Cai, J., Jepeal, L., Walton, D. S. and Maas, R. L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev* **8** (17), 2022-34.

Espindola, F. S., Suter, D. M., Partata, L. B., Cao, T., Wolenski, J. S., Cheney, R. E., King, S. M. and Mooseker, M. S. (2000). The light chain composition of chicken brain myosin-Va: calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN. *Cell Motil Cytoskeleton* **47** (4), 269-81.

Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M. and et al. (1995). Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet* **4** (3), 415-22.

Fantes, J., Ragge, N. K., Lynch, S. A., McGill, N. I., Collin, J. R., Howard-Peebles, P. N., Hayward, C., Vivian, A. J., Williamson, K., van Heyningen, V. and FitzPatrick, D. R. (2003). Mutations in SOX2 cause anophthalmia. *Nat Genet* **33** (4), 461-3.

Flint, J., Tufarelli, C., Peden, J., Clark, K., Daniels, R. J., Hardison, R., Miller, W., Philipsen, S., Tan-Un, K. C., McMorrow, T., Frampton, J., Alter, B. P., Frischauf, A. M. and Higgs, D. R. (2001). Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. *Hum Mol Genet* **10** (4), 371-82.

Foa, L., Rajan, I., Haas, K., Wu, G. Y., Brakeman, P., Worley, P. and Cline, H. (2001). The scaffold protein, Homer1b/c, regulates axon pathfinding in the central nervous system in vivo. *Nat Neurosci* **4** (5), 499-506.

Gehring, W. J. and Ikeo, K. (1999). Pax 6: mastering eye morphogenesis and eye evolution. *Trends Genet* **15** (9), 371-7.

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K.,

Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley Jr, R. L., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, J. and Rothberg, J. M. (2003). A Protein Interaction Map of Drosophila melanogaster. *Science*.

Glardon, S., Callaerts, P., Halder, G. and Gehring, W. J. (1997). Conservation of Pax-6 in a lower chordate, the ascidian Phallusia mammillata. *Development* **124** (4), 817-25.

Glaser, T., Walton, D. S. and Maas, R. L. (1992). Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat Genet* **2** (3), 232-9.

Glaser, T., Jepeal, L., Edwards, J. G., Young, S. R., Favor, J. and Maas, R. L. (1994). PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet* **7** (4), 463-71.

Gorlov, I. P. and Saunders, G. F. (2002). A method for isolating alternatively spliced isoforms: isolation of murine Pax6 isoforms. *Anal Biochem* **308** (2), 401-4.

Goudreau, G., Petrou, P., Reneker, L. W., Graw, J., Loster, J. and Gruss, P. (2002). Mutually regulated expression of Pax6 and Six3 and its implications for the Pax6 haploinsufficient lens phenotype. *Proc Natl Acad Sci U S A* **99** (13), 8719-24.

Grant, S. G. and O'Dell, T. J. (2001). Multiprotein complex signaling and the plasticity problem. *Curr Opin Neurobiol* **11** (3), 363-8.

Gronskov, K., Rosenberg, T., Sand, A. and Brondum-Nielsen, K. (1999). Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild aniridia phenotype. *Eur J Hum Genet* **7** (3), 274-86.

Gronskov, K., Olsen, J. H., Sand, A., Pedersen, W., Carlsen, N., Bak Jylling, A. M., Lyngbye, T., Brondum-Nielsen, K. and Rosenberg, T. (2001). Population-based risk estimates of Wilms tumor in sporadic aniridia. A comprehensive mutation screening procedure of PAX6 identifies 80% of mutations in aniridia. *Hum Genet* **109** (1), 11-8.

Halder, G., Callaerts, P. and Gehring, W. J. (1995). Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science* **267** (5205), 1788-92.

Hanson, I. and Van Heyningen, V. (1995). Pax6: more than meets the eye. *Trends Genet* **11** (7), 268-72.

Hanson, I., Churchill, A., Love, J., Axton, R., Moore, T., Clarke, M., Meire, F. and van Heyningen, V. (1999). Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations. *Hum Mol Genet* **8** (2), 165-72.

Hanson, I. M. (2003). PAX6 and congenital eye malformations. *Ped Res* **In Press**.

Heins, N., Malatesta, P., Cecconi, F., Nakafuku, M., Tucker, K. L., Hack, M. A., Chapouton, P., Barde, Y. A. and Gotz, M. (2002). Glial cells generate neurons: the role of the transcription factor Pax6. *Nat Neurosci* **5** (4), 308-15.

Hemmens, B., Woschitz, S., Pitters, E., Klosch, B., Volker, C., Schmidt, K. and Mayer, B. (1998). The protein inhibitor of neuronal nitric oxide synthase (PIN): characterization of its action on pure nitric oxide synthases. *FEBS Lett* **430** (3), 397-400.

Henry, J., Mather, I. H., McDermott, M. F. and Pontarotti, P. (1998). B30.2-like domain proteins: update and new insights into a rapidly expanding family of proteins. *Mol Biol Evol* **15** (12), 1696-705.

Herzig, R. P., Andersson, U. and Scarpulla, R. C. (2000). Dynein light chain interacts with NRF-1 and EWG, structurally and functionally related transcription factors from humans and drosophila. *J Cell Sci* **113 Pt 23**, 4263-73.

Heyman, I., Frampton, I., van Heyningen, V., Hanson, I., Teague, P., Taylor, A. and Simonoff, E. (1999). Psychiatric disorder and cognitive function in a family with an inherited novel mutation of the developmental control gene PAX6. *Psychiatr Genet* **9** (2), 85-90.

Hill, R. E., Favor, J., Hogan, B. L., Ton, C. C., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D. and van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354** (6354), 522-5.

Hodgson, S. V. and Saunders, K. E. (1980). A probable case of the homozygous condition of the aniridia gene. *J Med Genet* **17** (6), 478-80.

Hol, F. A., Geurds, M. P., Chatkupt, S., Shugart, Y. Y., Balling, R., Schrander-Stumpel, C. T., Johnson, W. G., Hamel, B. C. and Mariman, E. C. (1996). PAX genes and human neural tube defects: an amino acid substitution in PAX1 in a patient with spina bifida. *J Med Genet* **33** (8), 655-60.

Hoth, C. F., Milunsky, A., Lipsky, N., Sheffer, R., Clarren, S. K. and Baldwin, C. T. (1993). Mutations in the paired domain of the human PAX3 gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I). *Am J Hum Genet* **52** (3), 455-62.

Jacob, Y., Badrane, H., Ceccaldi, P. E. and Tordo, N. (2000). Cytoplasmic dynein LC8 interacts with lyssavirus phosphoprotein. *J Virol* **74** (21), 10217-22.

Jaffrey, S. R. and Snyder, S. H. (1996). PIN: an associated protein inhibitor of neuronal nitric oxide synthase. *Science* **274** (5288), 774-7.

Jans, D. A. and Hubner, S. (1996). Regulation of protein transport to the nucleus: central role of phosphorylation. *Physiol Rev* **76** (3), 651-85.

Johnson, M. and Lehtonen, J. (2000). Comparison of protein three-dimensional structures. *In Bioinformatics. Sequence, structure and databanks*. D. Higgins and W. Taylor. (Oxford University Press), 15-50.

Jordan, T., Hanson, I., Zaletayev, D., Hodgson, S., Prosser, J., Seawright, A., Hastie, N. and van Heyningen, V. (1992). The human PAX6 gene is mutated in two patients with aniridia. *Nat Genet* **1** (5), 328-32.

Kaiser, F. J., Tavassoli, K., Van den Bemd, G. J., Chang, G. T., Horsthemke, B., Moroy, T. and Ludecke, H. J. (2003). Nuclear interaction of the dynein light chain LC8a with the TRPS1

transcription factor suppresses the transcriptional repression activity of TRPS1. *Hum Mol Genet* **12** (11), 1349-58.

Kamachi, Y., Uchikawa, M. and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* **16** (4), 182-7.

Kamachi, Y., Uchikawa, M., Tanouchi, A., Sekido, R. and Kondoh, H. (2001). Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev* **15** (10), 1272-86.

Kammandel, B., Chowdhury, K., Stoykova, A., Aparicio, S., Brenner, S. and Gruss, P. (1999). Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. *Dev Biol* **205** (1), 79-97.

Kato, A., Ozawa, F., Saitoh, Y., Hirai, K. and Inokuchi, K. (1997). vesl, a gene encoding VASP/Ena family related protein, is upregulated during seizure, long-term potentiation and synaptogenesis. *FEBS Lett* **412** (1), 183-9.

Kato, A., Ozawa, F., Saitoh, Y., Fukazawa, Y., Sugiyama, H. and Inokuchi, K. (1998). Novel members of the Vesl/Homer family of PDZ proteins that bind metabotropic glutamate receptors. *J Biol Chem* **273** (37), 23969-75.

Keller, S. A., Jones, J. M., Boyle, A., Barrow, L. L., Killen, P. D., Green, D. G., Kapousta, N. V., Hitchcock, P. F., Swank, R. T. and Meisler, M. H. (1994). Kidney and retinal defects (Krd), a transgene-induced mutation with a deletion of mouse chromosome 19 that includes the Pax2 locus. *Genomics* **23** (2), 309-20.

King, S. M., Barbarese, E., Dillman, J. F., III, Patel-King, R. S., Carson, J. H. and Pfister, K. K. (1996). Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved Mr 8,000 light chain. *J Biol Chem* **271** (32), 19358-66.

Kleinjan, D. A., Seawright, A., Schedl, A., Quinlan, R. A., Danes, S. and van Heyningen, V. (2001). Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol Genet* **10** (19), 2049-59.

Kleinjan, D. A., Seawright, A., Elgar, G. and van Heyningen, V. (2002). Characterization of a novel gene adjacent to PAX6, revealing synteny conservation with functional significance. *Mamm Genome* **13** (2), 102-7.

Krawczak, M., Reiss, J. and Cooper, D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* **90** (1-2), 41-54.

Larsen, W. (1993). Development of the Eyes. *In Human Embryology*. (Churchill Livingstone Inc.), 341-351.

Lauderdale, J. D., Wilensky, J. S., Oliver, E. R., Walton, D. S. and Glaser, T. (2000). 3' deletions cause aniridia by preventing PAX6 gene expression. *Proc Natl Acad Sci U S A* **97** (25), 13755-9.

Macchia, P. E., Lapi, P., Krude, H., Pirro, M. T., Missero, C., Chiovato, L., Souabni, A., Baserga, M., Tassi, V., Pinchera, A., Fenzi, G., Gruters, A., Busslinger, M. and Di Lauro, R.

(1998). PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. *Nat Genet* **19** (1), 83-6.

Macdonald, R., Barth, K. A., Xu, Q., Holder, N., Mikkola, I. and Wilson, S. W. (1995). Midline signalling is required for Pax gene regulation and patterning of the eyes. *Development* **121** (10), 3267-78.

Mancino, F., Vekemans, M., Trasler, D. G. and Gros, P. (1992). Segregation analysis reveals tight genetic linkage between the spontaneously arising neural tube defect gene splotch (Sp) and Pax-3 in an intraspecific mouse backcross. *Cytogenet Cell Genet* **61** (2), 143-5.

Maquat, L. E. and Carmichael, G. G. (2001). Quality control of mRNA function. *Cell* **104** (2), 173-6.

Marquardt, T., Ashery-Padan, R., Andrejewski, N., Scardigli, R., Guillemot, F. and Gruss, P. (2001). Pax6 is required for the multipotent state of retinal progenitor cells. *Cell* **105** (1), 43-55.

Martha, A., Ferrell, R. E., Mintz-Hittner, H., Lyons, L. A. and Saunders, G. F. (1994). Paired box mutations in familial and sporadic aniridia predicts truncated aniridia proteins. *Am J Hum Genet* **54** (5), 801-11.

Mastick, G. S., Davis, N. M., Andrew, G. L. and Easter, S. S., Jr. (1997). Pax-6 functions in boundary formation and axon guidance in the embryonic mouse forebrain. *Development* **124** (10), 1985-97.

Meech, R., Kallunki, P., Edelman, G. M. and Jones, F. S. (1999). A binding site for homeodomain and Pax proteins is necessary for L1 cell adhesion molecule gene expression by Pax-6 and bone morphogenetic proteins. *Proc Natl Acad Sci U S A* **96** (5), 2420-5.

Mikkola, I., Bruun, J. A., Bjorkoy, G., Holm, T. and Johansen, T. (1999). Phosphorylation of the transactivation domain of Pax6 by extracellular signal-regulated kinase and p38 mitogen-activated protein kinase. *J Biol Chem* **274** (21), 15115-26.

Mishra, R., Gorlov, I. P., Chao, L. Y., Singh, S. and Saunders, G. F. (2002). PAX6, paired domain influences sequence recognition by the homeodomain. *J Biol Chem* **277** (51), 49488-94.

Mitchell, T. N., Free, S. L., Williamson, K. A., Stevens, J. M., Churchill, A. J., Hanson, I. M., Shorvon, S. D., Moore, A. T., van Heyningen, V. and Sisodiya, S. M. (2003). Polymicrogyria and absence of pineal gland due to PAX6 mutation. *Ann Neurol* **53** (5), 658-63.

Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O. (2001). Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* **8** (2), 85-95.

Naisbitt, S., Valtschanoff, J., Allison, D. W., Sala, C., Kim, E., Craig, A. M., Weinberg, R. J. and Sheng, M. (2000). Interaction of the postsynaptic density-95/guanylate kinase domain-associated protein complex with a light chain of myosin-V and dynein. *J Neurosci* **20** (12), 4524-34.

Neubuser, A., Koseki, H. and Balling, R. (1995). Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1. *Dev Biol* **170** (2), 701-16.

Niikura, T., Hashimoto, Y., Tajima, H., Ishizaka, M., Yamagishi, Y., Kawasumi, M., Nawa, M., Terashita, K., Aiso, S. and Nishimoto, I. (2003). A tripartite motif protein TRIM11 binds and destabilizes Humanin, a neuroprotective peptide against Alzheimer's disease-relevant insults. *Eur J Neurosci* **17** (6), 1150-8.

Nishina, S., Kohsaka, S., Yamaguchi, Y., Handa, H., Kawakami, A., Fujisawa, H. and Azuma, N. (1999). PAX6 expression in the developing human eye. *Br J Ophthalmol* **83** (6), 723-7.

Oikawa, T. and Yamada, T. (2003). Molecular biology of the Ets family of transcription factors. *Gene* **303**, 11-34.

Patel, P., Boyd, C. A., Johnston, D. G. and Williamson, C. (2002). Analysis of GAPDH as a standard for gene expression quantification in human placenta. *Placenta* **23** (8-9), 697-8.

Petralia, R. S., Wang, Y. X., Sans, N., Worley, P. F., Hammer, J. A., 3rd and Wenthold, R. J. (2001). Glutamate receptor targeting in the postsynaptic spine involves mechanisms that are independent of myosin Va. *Eur J Neurosci* **13** (9), 1722-32.

Planque, N., Leconte, L., Coquelle, F. M., Benkhelifa, S., Martin, P., Felder-Schmittbuhl, M. P. and Saule, S. (2001). Interaction of Maf transcription factors with Pax-6 results in synergistic activation of the glucagon promoter. *J Biol Chem* **276** (38), 35751-60.

Plaza, S., Dozier, C., Langlois, M. C. and Saule, S. (1995a). Identification and characterization of a neuroretina-specific enhancer element in the quail Pax-6 (Pax-QNR) gene. *Mol Cell Biol* **15** (2), 892-903.

Plaza, S., Dozier, C., Turque, N. and Saule, S. (1995b). Quail Pax-6 (Pax-QNR) mRNAs are expressed from two promoters used differentially during retina development and neuronal differentiation. *Mol Cell Biol* **15** (6), 3344-53.

Poisson, N., Real, E., Gaudin, Y., Vaney, M. C., King, S., Jacob, Y., Tordo, N. and Blondel, D. (2001). Molecular basis for the interaction between rabies virus phosphoprotein P and the dynein light chain LC8: dissociation of dynein-binding properties and transcriptional functionality of P. *J Gen Virol* **82** (Pt 11), 2691-6.

Ponting, C., Schultz, J. and Bork, P. (1997). SPRY domains in ryanodine receptors (Ca(2+)-release channels). *Trends Biochem Sci* **22** (6), 193-4.

Prosser, J. and van Heyningen, V. (1998). PAX6 mutations reviewed. *Hum Mutat* **11** (2), 93-108.

Puthalakath, H., Huang, D. C., O'Reilly, L. A., King, S. M. and Strasser, A. (1999). The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex. *Mol Cell* **3** (3), 287-96.

Quiring, R., Walldorf, U., Kloter, U. and Gehring, W. J. (1994). Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. *Science* **265** (5173), 785-9.

Rakic, P. and Lombroso, P. J. (1998). Development of the cerebral cortex: I. Forming the cortical structure. *J Am Acad Child Adolesc Psychiatry* **37** (1), 116-7.

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., Guffanti, A., Minucci, S., Pelicci, P. G. and Ballabio, A. (2001). The tripartite motif family identifies cell compartments. *Embo J* **20** (9), 2140-51.

Rezvani, K., Mee, M., Dawson, S., McIlhinney, J., Fujita, J. and Mayer, R. J. (2003). Proteasomal interactors control activities as diverse as the cell cycle and glutaminergic neurotransmission. *Biochem Soc Trans* **31** (2), 470-3.

Rodriguez-Crespo, I., Straub, W., Gavilanes, F. and Ortiz de Montellano, P. R. (1998). Binding of dynein light chain (PIN) to neuronal nitric oxide synthase in the absence of inhibition. *Arch Biochem Biophys* **359** (2), 297-304.

Rodriguez-Crespo, I., Yelamos, B., Roncal, F., Albar, J. P., Ortiz de Montellano, P. R. and Gavilanes, F. (2001). Identification of novel cellular proteins that bind to the LC8 dynein light chain using a pepscan technique. *FEBS Lett* **503** (2-3), 135-41.

Rong, R., Ahn, J. Y., Huang, H., Nagata, E., Kalman, D., Kapp, J. A., Tu, J., Worley, P. F., Snyder, S. H. and Ye, K. (2003). PI3 kinase enhancer-Homer complex couples mGluRI to PI3 kinase, preventing neuronal apoptosis. *Nat Neurosci* **6** (11), 1153-61.

Sakai, M., Serria, M. S., Ikeda, H., Yoshida, K., Imaki, J. and Nishi, S. (2001). Regulation of c-maf gene expression by Pax6 in cultured cells. *Nucleic Acids Res* **29** (5), 1228-37.

Sala, C., Piech, V., Wilson, N. R., Passafaro, M., Liu, G. and Sheng, M. (2001). Regulation of dendritic spine morphology and synaptic function by Shank and Homer. *Neuron* **31** (1), 115-30.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74** (12), 5463-7.

Saurin, A. J., Borden, K. L., Boddy, M. N. and Freemont, P. S. (1996). Does this have a familiar RING? *Trends Biochem Sci* **21** (6), 208-14.

Schwarz, M., Cecconi, F., Bernier, G., Andrejewski, N., Kammandel, B., Wagner, M. and Gruss, P. (2000). Spatial specification of mammalian eye territories by reciprocal transcriptional repression of Pax2 and Pax6. *Development* **127** (20), 4325-34.

Sheng, M. and Kim, M. J. (2002). Postsynaptic signaling and plasticity mechanisms. *Science* **298** (5594), 776-80.

Shin, D. H., Kwon, B. S., Chang, Y. P., Bae, S. R., Kim, J. and Kim, J. W. (2003). Ultramicroscopical immunolocalization of PAX6 in the adult chicken retina. *Acta Histochem* **105** (3), 267-72.

Shiraishi, Y., Mizutani, A., Bito, H., Fujisawa, K., Narumiya, S., Mikoshiba, K. and Furuichi, T. (1999). Cupidin, an isoform of Homer/Vesl, interacts with the actin cytoskeleton and activated rho family small GTPases and is expressed in developing mouse cerebellar granule cells. *J Neurosci* **19** (19), 8389-400.

Shiraishi, Y., Mizutani, A., Mikoshiba, K. and Furuichi, T. (2003). Coincidence in dendritic clustering and synaptic targeting of homer proteins and NMDA receptor complex proteins NR2B and PSD95 during development of cultured hippocampal neurons. *Mol Cell Neurosci* **22** (2), 188-201.

Singh, S., Tang, H. K., Lee, J. Y. and Saunders, G. F. (1998). Truncation mutations in the transactivation region of PAX6 result in dominant-negative mutants. *J Biol Chem* **273** (34), 21531-41.

Singh, S., Chao, L. Y., Mishra, R., Davies, J. and Saunders, G. F. (2001). Missense mutation at the C-terminus of PAX6 negatively modulates homeodomain function. *Hum Mol Genet* **10** (9), 911-8.

Singh, S., Mishra, R., Arango, N. A., Deng, J. M., Behringer, R. R. and Saunders, G. F. (2002). Iris hypoplasia in mice that lack the alternatively spliced Pax6(5a) isoform. *Proc Natl Acad Sci U S A* **99** (10), 6812-5.

Sisodiya, S. M., Free, S. L., Williamson, K. A., Mitchell, T. N., Willis, C., Stevens, J. M., Kendall, B. E., Shorvon, S. D., Hanson, I. M., Moore, A. T. and van Heyningen, V. (2001). PAX6 haploinsufficiency causes cerebral malformation and olfactory dysfunction in humans. *Nat Genet* **28** (3), 214-6.

Skurat, A. V., Dietrich, A. D., Zhai, L. and Roach, P. J. (2002). GNIP, a novel protein that binds and activates glycogenin, the self-glucosylating initiator of glycogen biosynthesis. *J Biol Chem* **277** (22), 19331-8.

Stockton, D. W., Das, P., Goldenberg, M., D'Souza, R. N. and Patel, P. I. (2000). Mutation of PAX9 is associated with oligodontia. *Nat Genet* **24** (1), 18-9.

Stoykova, A., Fritsch, R., Walther, C. and Gruss, P. (1996). Forebrain patterning defects in Small eye mutant mice. *Development* **122** (11), 3453-65.

Sun, J., Tadokoro, S., Imanaka, T., Murakami, S. D., Nakamura, M., Kashiwada, K., Ko, J., Nishida, W. and Sobue, K. (1998). Isolation of PSD-Zip45, a novel Homer/vesl family protein containing leucine zipper motifs, from rat brain. *FEBS Lett* **437** (3), 304-8.

Tajima, H., Niikura, T., Hashimoto, Y., Ito, Y., Kita, Y., Terashita, K., Yamazaki, K., Koto, A., Aiso, S. and Nishimoto, I. (2002). Evidence for in vivo production of Humanin peptide, a neuroprotective factor against Alzheimer's disease-related insults. *Neurosci Lett* **324** (3), 227-31.

Takahashi, M. and Osumi, N. (2002). Pax6 regulates specification of ventral neurone subtypes in the hindbrain by establishing progenitor domains. *Development* **129** (6), 1327-38.

Talamillo, A., Quinn, J. C., Collinson, J. M., Caric, D., Price, D. J., West, J. D. and Hill, R. E. (2003). Pax6 regulates regional development and neuronal migration in the cerebral cortex. *Dev Biol* **255** (1), 151-63.

Tang, H. K., Singh, S. and Saunders, G. F. (1998). Dissection of the transactivation function of the transcription factor encoded by the eye developmental gene PAX6. *J Biol Chem* **273** (13), 7210-21.

Thomas, U. (2002). Modulation of synaptic signalling complexes by Homer proteins. *J Neurochem* **81** (3), 407-13.

Tomarev, S. I., Callaerts, P., Kos, L., Zinovieva, R., Halder, G., Gehring, W. and Piatigorsky, J. (1997). Squid Pax-6 and eye development. *Proc Natl Acad Sci U S A* **94** (6), 2421-6.

Tominaga, T., Meng, W., Togashi, K., Urano, H., Alberts, A. S. and Tominaga, M. (2002). The Rho GTPase effector protein, mDia, inhibits the DNA binding ability of the transcription factor Pax6 and changes the pattern of neurite extension in cerebellar granule cells through its binding to Pax6. *J Biol Chem* **277** (49), 47686-91.

Ton, C. C., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., van Heyningen, V., Hastie, N. D., Meijers-Heijboer, H., Drechsler, M. and et al. (1991). Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* **67** (6), 1059-74.

Topcu, Z., Mack, D. L., Hromas, R. A. and Borden, K. L. (1999). The promyelocytic leukemia protein PML interacts with the proline-rich homeodomain protein PRH: a RING may link hematopoiesis and growth control. *Oncogene* **18** (50), 7091-100.

Tu, J. C., Xiao, B., Yuan, J. P., Lanahan, A. A., Leoffert, K., Li, M., Linden, D. J. and Worley, P. F. (1998). Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. *Neuron* **21** (4), 717-26.

Tu, J. C., Xiao, B., Naisbitt, S., Yuan, J. P., Petralia, R. S., Brakeman, P., Doan, A., Aakalu, V. K., Lanahan, A. A., Sheng, M. and Worley, P. F. (1999). Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins. *Neuron* **23** (3), 583-92.

Vale, R. D. (2003). The molecular motor toolbox for intracellular transport. *Cell* **112** (4), 467-80.

Vallee, R. B., Tai, C. and Faulkner, N. E. (2001). LIS1: cellular function of a disease-causing gene. *Trends Cell Biol* **11** (4), 155-60.

Vincent, M. C., Pujo, A. L., Olivier, D. and Calvas, P. (2003). Screening for PAX6 gene mutations is consistent with haploinsufficiency as the main mechanism leading to various ocular defects. *Eur J Hum Genet* **11** (2), 163-9.

Walther, C. and Gruss, P. (1991). Pax-6, a murine paired box gene, is expressed in the developing CNS. *Development* **113** (4), 1435-49.

Westerman, B. A., Murre, C. and Oudejans, C. B. (2003). The cellular Pax-Hox-helix connection. *Biochim Biophys Acta* **1629** (1-3), 1-7.

Whang-Peng, J., Knutsen, T., Theil, K., Horowitz, M. E. and Triche, T. (1992). Cytogenetic studies in subgroups of rhabdomyosarcoma. *Genes Chromosomes Cancer* **5** (4), 299-310.

Williams, S. C., Altmann, C. R., Chow, R. L., Hemmati-Brivanlou, A. and Lang, R. A. (1998). A highly conserved lens transcriptional control element from the Pax-6 gene. *Mech Dev* **73** (2), 225-9.

Wilson, D. S., Guenther, B., Desplan, C. and Kuriyan, J. (1995). High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell* **82** (5), 709-19.

Wilson, M. J., Salata, M. W., Susalka, S. J. and Pfister, K. K. (2001). Light chains of mammalian cytoplasmic dynein: identification and characterization of a family of LC8 light chains. *Cell Motil Cytoskeleton* **49** (4), 229-40.

Xiao, B., Tu, J. C., Petralia, R. S., Yuan, J. P., Doan, A., Breder, C. D., Ruggiero, A., Lanahan, A. A., Wenthold, R. J. and Worley, P. F. (1998). Homer regulates the association of group 1 metabotropic glutamate receptors with multivalent complexes of homer-related, synaptic proteins. *Neuron* **21** (4), 707-16.

Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L. and Pabo, C. O. (1999a). Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev* **13** (10), 1263-75.

Xu, P. X., Zhang, X., Heaney, S., Yoon, A., Michelson, A. M. and Maas, R. L. (1999b). Regulation of Pax6 expression is conserved between mice and flies. *Development* **126** (2), 383-95.

Xu, W., Rould, M. A., Jun, S., Desplan, C. and Pabo, C. O. (1995). Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. *Cell* **80** (4), 639-50.

Yamasaki, T., Kawaji, K., Ono, K., Bito, H., Hirano, T., Osumi, N. and Kengaku, M. (2001). Pax6 regulates granule cell polarization during parallel fiber formation in the developing cerebellum. *Development* **128** (16), 3133-44.

Zahraoui, A., Touchot, N., Chardin, P. and Tavitian, A. (1989). The human Rab genes encode a family of GTP-binding proteins related to yeast YPT1 and SEC4 products involved in secretion. *J Biol Chem* **264** (21), 12394-401.

Zhang, X., Friedman, A., Heaney, S., Purcell, P. and Maas, R. L. (2002). Meis homeoproteins directly regulate Pax6 during vertebrate lens morphogenesis. *Genes Dev* **16** (16), 2097-107.

Zhang, Y. and Emmons, S. W. (1995). Specification of sense-organ identity by a Caenorhabditis elegans Pax-6 homologue. *Nature* **377** (6544), 55-9.

Zhang, Y., Ferreira, H. B., Greenstein, D., Chisholm, A. and Emmons, S. W. (1998). Regulated nuclear entry of the C. elegans Pax-6 transcription factor. *Mech Dev* **78** (1-2), 179-87.

Zhou, Y., Zheng, J. B., Gu, X., Li, W. and Saunders, G. F. (2000). A novel Pax-6 binding site in rodent B1 repetitive elements: coevolution between developmental regulation and repeated elements? *Gene* **245** (2), 319-28.

Zhou, Y. H., Zheng, J. B., Gu, X., Saunders, G. F. and Yung, W. K. (2002). Novel PAX6 binding sites in the human genome and the role of repetitive elements in the evolution of gene regulation. *Genome Res* **12** (11), 1716-22.